



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 94276

TO: Manjunath N Rao
Location: CM1/10D01/10A11
Art Unit: 1652
Wednesday, May 28, 2003

Case Serial Number: 040863

From: Susan Hanley
Location: Biotech-Chem Library
CM1-6B05
Phone: 305-4053

susan.hanley@uspto.gov

Search Notes

10/040863

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 ; Search time 37.6369 seconds
(without alignments)
1932.533 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKQLPLSEKELPMTT.....RSFHLKAKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1896	100.0	353	11 055025	O55025 rattus norv
2	1896	100.0	380	11 09JK44	Q9JK44 rattus norv
3	1747	92.1	354	11 035087	Q35087 rattus norv
4	1747	92.1	354	11 09R275	O9R275 rattus norv
5	1538.5	81.1	347	11 09JL27	Q9JL27 mus musculus
6	1534.5	80.9	347	11 070504	O70504 mus musculus
7	1483.5	78.2	321	11 0920W2	Q920W2 mus spicile
8	1483.5	78.2	321	11 091VF0	Q91VF0 mus musculus
9	1468.5	77.5	321	11 0920W3	Q920W3 mus musculus
10	1400.5	73.9	343	6 077486	O77486 gorilla gor
11	1397.5	73.7	338	6 09TUD5	O9TUD5 pan troglod
12	1397.5	73.7	343	6 077485	O77485 pan troglod
13	1396.5	73.7	343	6 09TUD3	Q9TUD3 gorilla gor
14	1391.5	73.4	347	6 029505	Q29505 oryctolagus
15	1389.5	73.3	343	6 077487	O77487 pongo pygma
16	1382.5	72.9	332	4 099450	O99450 homo sapien

17	1382.5	72.9	343	6 077712	O77712 cercopithec
18	1379.5	72.8	343	6 09TTC7	Q9TTC7 hylobates i
19	1346	71.0	368	11 0920T3	Q920T3 mus musculus
20	1343	70.8	299	11 09WCE6	Q9WCE6 rattus norv
21	1342	70.8	368	11 0920T2	Q920T2 mus spicile
22	1339	70.6	346	4 014338	Q14338 homo sapien
23	1338	70.6	368	11 0920T4	Q920T4 mus musculus
24	1330	70.1	368	11 091V73	Q91V73 mus musculus
25	1319	69.6	273	6 09TSL5	O9TSL5 macaca fasc
26	1307	68.9	273	6 09TTC9	Q9TTC9 eulemur ful
27	1262	66.6	273	6 09TTD0	Q9TTD0 callithrix
28	1229	64.8	348	6 09TUD1	Q9TUD1 macaca mula
29	1214	64.0	368	6 09TTY3	Q9TTY3 bos taurus
30	1207	63.7	332	6 09N266	Q9N266 macaca fasc
31	1109	58.5	377	11 09JL28	O9JL28 mus musculus
32	1109	58.5	377	11 09T327	P9T327 mus musculus
33	1099	58.0	377	11 0920W5	Q920W5 mus musculus
34	1099	58.0	377	11 0920W4	Q920W4 mus spicile
35	1086.5	57.3	360	6 09TTY7	Q9TTY7 bos taurus
36	1051.5	55.5	366	6 09TUD4	O9TUD4 gorilla gor
37	1050.5	55.4	366	6 09TUE6	O9TUE6 eulemur ful
38	1048.5	55.3	365	6 09TUE7	O9TUE7 hylobates i
39	1048	55.3	366	6 09TUD2	Q9TUD2 macaca mula
40	1047.5	55.2	353	6 09N270	Q9N270 pongo pygma
41	1047.5	55.2	365	4 08WXB6	O8WXB6 homo sapien
42	1047	55.2	366	6 09N267	O9N267 macaca fasc
43	1046.5	55.2	365	4 014505	O14505 homo sapien
44	1043.5	55.0	366	6 09TUD6	O9TUD6 pan troglod
45	1042.5	55.0	365	4 014506	O14506 homo sapien

ALIGNMENTS

RESULT 1

ID: O55025 PRELIMINARY: PRT: 353 AA.

AC O55025;
DT 01-JUN-1998 (TRENBLrei. 06, Created)
DT 01-JUN-1998 (TRENBLrei. 06, Last sequence update)
DE 01-MAR-2002 (TRENBLrei. 20, Last annotation update)
DE Alpha 1,2 fucosyltransferase (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BUFFALO RAT;
RX MEDLINE=98342056; PubMed=9675030;
RA Sherwood A.L., Holmes E.H.;
RT "Cloning and expression of the catalytic domain from rat hepatoma H35
RT cell GDP-fucose: GM1 alpha 1-->2fucosyltransferase, an enzyme which is
RT activated during early stages of chemical carcinogenesis in rat
RT liver".
RL Arch. Biochem. Biophys. 355:215-221(1998).
DR EMBL; AF042743; AAC14695.1; .
DR InterPro; IPR002516; GL11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 353 AA; 39738 MW; 609D64EB222C9585 CRC64;

Query Match 100.0%; Score 1896; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.3e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKQLPLSEKELPMTTMSNGTSPENRRDSEQHNGELRCGMFTINSIGRLGNQM 60
|||||
DB 1 LQQRIVKQLPLSEKELPMTTMSNGTSPENRRDSEQHNGELRCGMFTINSIGRLGNQM 60
|||||
QY 61 GYATLAFALARNGLRIATPASHMNALAFIFRISLPVLSHTAKKIPWQNYHLNDWMEER 120
|||||


```
Db 61 GEYATLFLALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
QY 121 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 121 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
QY 181 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 240
Db 181 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 240
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 300
Db 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 300
QY 301 LKVFPEAAFLPEWVGIPADLSPLLKALIPACPRSHFHLKAGVTCYVAGRAF 353
Db 301 LKVFPEAAFLPEWVGIPADLSPLLKALIPACPRSHFHLKAGVTCYVAGRAF 353
RESULT 2
Q5UK44 PRELIMINARY; PRT: 380 AA.
ID Q5JK44
AC Q5JK44;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1-2 fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER;
RA Sherwood A.L.; Stroud M.R.; Levery S.B.; Holmes E.H.;
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-
RT fucose:GM1 alpha 1-2 fucosyltransferase is required for optimum enzyme
RT activity and interaction with lipids";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264005; AAF72200.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;
Query Match 100.0%; Score 1896; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 4, 7e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQORIVKLOPLSEKELPMTTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNOM 60
Db 28 LQORIVKLOPLSEKELPMTTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNOM 87
QY 61 GEYATLFLALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 88 GEYATLFLALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 147
QY 121 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 148 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
QY 181 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 240
Db 208 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 267
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 300
Db 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 327
QY 301 LKVFPEAAFLPEWVGIPADLSPLLK 326
Db 328 LKVFPEAAFLPEWVGIPADLSPLLK 353
RESULT 4
Q9R275 PRELIMINARY; PRT: 354 AA.
ID Q9R275
AC Q9R275;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=21099374; PubMed=11179967;
```

RESULT 3

```
Q35087 PRELIMINARY; PRT: 354 AA.
ID Q35087
AC Q35087;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON CANCER;
RA Soejima M.; Wang B.; Koda Y.; Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RT fucosyltransferase genes";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006138; BAA21742.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 8636444888215BA1 CRC64;
Query Match 92.1%; Score 1747; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e-150;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LQORIVKLOPLSEKELPMTTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNOM 60
Db 28 LQORIVKLOPLSEKELPMTTQMSSGNTESPMRRDSEQHNGELRGMTINSIGRLGNOM 87
QY 61 GEYATLFLALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 120
Db 88 GEYATLFLALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER 147
QY 121 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
Db 148 YRHIPGHEVFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
QY 181 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 240
Db 208 GYHVVRGDIYHVMNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMWCARENINAS 267
QY 241 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 300
Db 268 RGVVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLANITLPDSPF 327
QY 301 LKVFPEAAFLPEWVGIPADLSPLLK 326
Db 328 LKVFPEAAFLPEWVGIPADLSPLLK 353
RESULT 4
Q9R275 PRELIMINARY; PRT: 354 AA.
ID Q9R275
AC Q9R275;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=21099374; PubMed=11179967;
```

```
RA Bureau V., Marionneau S., Caillaud-Thomas A., Le Moullac-Vaidye B.,
RA Liehr T., Le Pendu J.,
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL: AF131238; AAD24469.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase
SQ SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;

Query Match 92.1%; Score 1747; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e-150;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTTQSSNGNTSEPMRRDSEHQNGELRGMTFINSIGRLGNQM 60
DB 28 LQORIVKLOPLSEKEL-QAVQSSPNAARTDMQ-----SAKLOGIFTINSIGRLGNQM 80
QY 61 GEYATLAFALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWONYHLNDWMEER 120
DB 81 GEYATLAFALARMNGRLAFIPESMHNALAFIRISLPVLSHSDTARRIPWONYHLNDWMEER 140
QY 121 YRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 180
DB 141 YRHIPQYVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 200
QY 181 GVHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSPVFTVTSNGMAWCRENINAS 240
DB 201 GVHVRGDYVHVMPKVKGVADRGYLEKALDRFRARYSPVFTVTSNGMAWCRENINTS 260
QY 241 RGDVVPAGNGIEGSPAKDFALLTQCNTMTICTFGIWAAYLAGGDTIYLANLYLPDSPF 300
DB 261 LGDVVPAGNGIEGSPAKDFALLTQCNTMTICTFGIWAAYLAGGDTIYLANLYLPDSPF 320
QY 301 LKVFKEPAAFLPEWVGIPADLSPLLK 326
DB 321 LKTFKPAARFLPEWVGIPADLSPLLK 346

RESULT 6
O70504 PRELIMINARY; PRT; 347 AA.
ID O70504
AC O70504;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
DE FUT2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RT "GDP-fucose: beta-galactoside alphaL,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide.";
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL: AF064792; AAC16887.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;

Query Match 80.9%; Score 1534.5; DB 11; Length 347;
Best Local Similarity 88.7%; Pred. No. 3.1e-131;
Matches 289; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQORIVKLOPLSEKELPMTTQSSNGNTSEPMRRDSEHQNGELRGMTFINSIGRLGNQM 60
DB 28 LQORIVKLOPLSEKEL-QAVQSSPNAARTDMQ-----SAKLOGIFTINSIGRLGNQM 80
QY 61 GEYATLAFALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWONYHLNDWMEER 120
DB 81 GEYATLAFALARMNGRLAFIPESMHNALAFIRISLPVLSHSDTARRIPWONYHLNDWMEER 140
QY 121 YRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 180
DB 141 YRHIPQYVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 200

RA Bureau V., Marionneau S., Caillaud-Thomas A., Le Moullac-Vaidye B.,
RA Liehr T., Le Pendu J.,
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL: AF131238; AAD24469.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase
SQ SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;

Query Match 92.1%; Score 1747; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.5e-150;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQORIVKLOPLSEKELPMTTQSSNGNTSEPMRRDSEHQNGELRGMTFINSIGRLGNQM 60
DB 28 LQORIVKLOPLSEKELPMTTQSSNGNTSEPMRRDSEHQNGELRGMTFINSIGRLGNQM 87
QY 61 GEYATLAFALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWONYHLNDWMEER 120
DB 81 GEYATLAFALARMNGRLAFIPASMHNALAFIRISLPVLSHSDTAKKIPWONYHLNDWMEER 147
QY 121 YRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 180
DB 148 YRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPSTFV 207
QY 181 GVHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSPVFTVTSNGMAWCRENINAS 240
DB 208 GVHVRGDYVHVMPNVKGVADRGYLEKALDMFRARYSPVFTVTSNGMAWCRENINAS 267
QY 241 RGDVVPAGNGIEGSPAKDFALLTQCNTMTICTFGIWAAYLAGGDTIYLANLYLPDSPF 300
DB 268 RGDVVPAGNGIEGSPAKDFALLTQCNTMTICTFGIWAAYLAGGDTIYLANLYLPDSPF 327
QY 301 LKVFKEPAAFLPEWVGIPADLSPLLK 326
DB 328 LKVFKEPAAFLPEWVGIPADLSPLLK 353

RESULT 5
O9JL27 PRELIMINARY; PRT; 347 AA.
ID O9JL27
AC O9JL27
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
DE FUT2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL: AF214656; AAF45146.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046C8E79 CRC64;

Query Match 81.1%; Score 1538.5; DB 11; Length 347;
Best Local Similarity 89.0%; Pred. No. 1.4e-131;
Matches 290; Conservative 12; Mismatches 17; Indels 7; Gaps 2;
```

```

QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 201 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINTS 260
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 300
Db 261 LGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 320
QY 301 LKVFPEAAFLPEWVGPADLSPLLK 326
Db 321 LKIFKPAFAFLPEWVGPADLSPLLK 346

RESULT 7
Q920W2 PRELIMINARY; PRT; 321 AA.
AC Q920W2
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koida T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039123; BAB68647.1; -
DR EMBL: AB039123; BAB68647.1; -
DR InterPro: IPR002516; GT_11; -
DR Pfam: PF01531; Glyco_transferase.
KW Glycosyltransferase; transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36560 MW; 5B7D19BA6D4B5394 CRC64;

```

```

Query Match 78.2%; Score 1483.5; DB 11; Length 321;
Best Local Similarity 88.9%; Pred. No. 1.2e-126;
Matches 280; Conservative 11; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQORIVKLOPSEKELPMTQMSSGNTSPERMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 14 LQORIVKLOPSEKEL-QAIQSSLSNAARTMQQ-----SAKLOGITINSIGRLGNQM 66

QY 61 GEYATLALARMNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPWONYHLNDWMEER 120
Db 67 GEYATLALARMNGRLATIPESMHNALAPIFRISLPVLHSDTARRIPWONYHLNDWMEER 126

QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSOPSTFV 180
Db 127 YRHIPGVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSOPSTFV 186

QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 187 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINTS 246

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 300
Db 247 LGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 306

QY 301 LKVFPEAAFLPEWV 315
Db 307 LKIFKPAFAFLPEW 321

```

```

RESULT 8
Q91VF0 PRELIMINARY; PRT; 321 AA.
AC Q91VF0
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.

```

```

ID Q91VF0 PRELIMINARY; PRT; 321 AA.
AC Q91VF0
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Liu Y., Kitano T., Koida T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039114; BAB68638.1; -
DR EMBL: AB039115; BAB68639.1; -
DR EMBL: AB039116; BAB68640.1; -
DR EMBL: AB039117; BAB68641.1; -
DR EMBL: AB039118; BAB68642.1; -
DR EMBL: AB039119; BAB68643.1; -
DR EMBL: AB039120; BAB68644.1; -
DR EMBL: AB039121; BAB68645.1; -
DR InterPro: IPR002516; GT_11; -
DR Pfam: PF01531; Glyco_transferase.
KW Glycosyltransferase; transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36516 MW; 0622D3BB503B72D1 CRC64;

```

```

Query Match 78.2%; Score 1483.5; DB 11; Length 321;
Best Local Similarity 88.6%; Pred. No. 1.2e-126;
Matches 279; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

QY 1 LQORIVKLOPSEKELPMTQMSSGNTSPERMRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 14 LQORIVKLOPSEKEL-QAVQWSSPNAARTMQQ-----SAKLOGITINSIGRLGNQM 66

QY 61 GEYATLALARMNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPWONYHLNDWMEER 120
Db 67 GEYATLALARMNGRLATIPESMHNALAPIFRISLPVLHSDTARRIPWONYHLNDWMEER 126

QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSOPSTFV 180
Db 127 YRHIPGVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVEEAQAFRLGLRVNGSOPSTFV 186

QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 187 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINTS 246

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 300
Db 247 LGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTGTGFIWAAYLAGGDTIYLANITLPDPSF 306

QY 301 LKVFPEAAFLPEWV 315
Db 307 LKIFKPAFAFLPEW 321

```

```

RESULT 9
Q920W3 PRELIMINARY; PRT; 321 AA.
AC Q920W3
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC 2.4.1.69) (Fragment).
GN FUT2.

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NJL/MSF;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039122; BAB69646.1; -
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 321
FT NON_TER 321 321
SQ SEQUENCE 321 AA; 36464 MW; FF5304CD150F774A CRC64;

Query Match 77.5%; Score 1468.5; DB 11; Length 321;
Best Local Similarity 87.9%; Pred. No. 2.8e-125;
Matches 277; Conservative 13; Mismatches 18; Indels 7; Gaps 2;

QY 1 LQORIVKQLPSLSEKELPMTQMSSGNTEPEMRDSEQHNGELRGMFTINSIGRLGNQM 60
Db 14 LQOQIVKQLPSLSEKELPMTQMSSGNTEPEMRDSEQHNGELRGMFTINSIGRLGNQM 66
QY 61 GEYATLALARMNGRLAFIPASHMNLALAFIRISLPVHSDTAKKIPWQNYHLNDWMEER 120
Db 67 GEYATLALARMNGRLAFIPASHMNLALAFIRISLPVHSDTAKKIPWQNYHLNDWMEER 126
QY 121 YRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGLVNGSQPSTFV 180
Db 127 YRHIPGVYVFTGYPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGLVNGSQPSTFV 186
QY 181 GVHVRGDYVHVMPKVKVGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENINAS 240
Db 187 GVHVRGDYVHVMPKVKVGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENINAS 246
QY 241 RGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGGDTIYLYANTLPDPSF 300
Db 247 LGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGGDTIYLYANTLPDPSF 306
QY 301 LKVFKEPAEAFLEPWV 315
Db 307 LKIFKPAEAFLEPWV 321

RESULT 10
O77486 PRELIMINARY; PRT; 343 AA.
AC O77486;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-Mar-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele s428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";

RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015635; BAA31128.1; -
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FBC CRC64;

Query Match 73.9%; Score 1400.5; DB 6; Length 343;
Best Local Similarity 79.8%; Pred. No. 4.8e-119;
Matches 260; Conservative 25; Mismatches 30; Indels 11; Gaps 2;

QY 1 LQORIVKQLPSLSEKELPMTQMSSGNTEPEMRDSEQHNGELRGMFTINSIGRLGNQM 60
Db 28 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMWTINAGRLGNQM 76
QY 61 GEYATLALARMNGRLAFIPASHMNLALAFIRISLPVHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLALAKMNGRPAFIPAQNHSTLAPIFRTLPVHLSATASRIQPWNYHLNDWMEER 136
QY 121 YRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGLVNGSQPSTFV 180
Db 137 YRHIPGVYVFTGYPCSWTFYHHLRPEILKEFTLHDVREEAQAFRLGLVNGSQPSTFV 196
QY 181 GVHVRGDYVHVMPKVKVGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENINAS 240
Db 197 GVHVRGDYVHVMPKVKVGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCRENIDTS 256
QY 241 RGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGGDTIYLYANTLPDPSF 300
Db 257 HGDVVFAGDIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGGDTIYLYANTLPDPSF 316
QY 301 LKVFKEPAEAFLEPWVGPADLSPLK 326
Db 317 LKIFKPAEAFLEPWGTGAADLSPLK 342

RESULT 11
Q9TUD5 PRELIMINARY; PRT; 338 AA.
AC Q9TUD5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-Mar-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyl transferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH16;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: re-ation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080604; AAF14066.1; -
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 338 AA; 38428 MW; E7FABA0FF1BC95F9 CRC64;

Query Match 73.7%; Score 1397.5; DB 6; Length 338;
Best Local Similarity 79.4%; Pred. No. 8.7e-119;
Matches 259; Conservative 27; Mismatches 29; Indels 11; Gaps 2;

QY 1 LQORIVKQLPSLSEKELPMTQMSSGNTEPEMRDSEQHNGELRGMFTINSIGRLGNQM 60
Db 23 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMWTINAGRLGNQM 71

```
Oy 61 GEYATLFALARMNGRLAFIPASMHNALAFIRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 72 GEYATLYALAKMNGRPAFIPAQMHSTLAFIRITLPVLHSATASRIPWQNYHLNDWMEEE 131
Oy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREEAQAFLRGLRVNGSPSTFV 180
Db 132 YRHIPGEVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREEAQAFLRGLRVNGSRPGTFV 191
Oy 181 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 192 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENIDTS 251
Oy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGGDTIYLANYLTPDSPF 300
Db 252 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGGDTIYLANYLTPDSPF 311
Oy 301 LKVKPEAAFLPEWVGIPADLSPLK 326
Db 312 LKIFKPEAAFLPENMGIAADLSPLK 337

RESULT 12
O77485 PRELIMINARY: PRT: 343 AA.
ID O77485
AC O77485;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DE 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RL alpha(1,2) fucosyltransferase gene (FUT2).";
RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RL group alpha(1,2) fucosyltransferase gene (FUT2).";
RN J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015634; BAA31127.1; -.
DR InterPro; IPR002516; GP_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Glyco-transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match 73.78; Score 1397.5; DB 6; Length 343;
Best Local Similarity 79.4%; Pred. No. 8.9e-119;
Matches 259; Conservative 27; Mismatches 29; Indels 11; Gaps 2;

Oy 1 LQORIVKLQPLSEKELPMTTQSSNTSPESMRDSEOHNGELRGMTINSIGRLGNQM 60
Db 28 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGRLGNQM 76
Oy 61 GEYATLFALARMNGRLAFIPASMHNALAFIRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLYALAKMNGRPAFIPAQMHSTLAFIRITLPVLHSATASRIPWQNYHLNDWMEEE 136
Oy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREEAQAFLRGLRVNGSPSTFV 180
Db 137 YRHIPGEVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREEAQAFLRGLRVNGSRPGTFV 196
Oy 181 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 192 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENIDTS 256
Oy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGGDTIYLANYLTPDSPF 300
Db 252 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGGDTIYLANYLTPDSPF 316
Oy 301 LKVKPEAAFLPEWVGIPADLSPLK 326
Db 317 LKIFKPEAAFLPENMGIAADLSPLK 342

RESULT 14
Q29505 PRELIMINARY: PRT: 347 AA.
ID Q29505
AC Q29505;
DT 01-FEB-1997 (TremBLrel. 02, Created)
DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
```

```
Oy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGGDTIYLANYLTPDSPF 300
Db 257 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGGDTIYLANYLTPDSPF 316
Oy 301 LKVKPEAAFLPEWVGIPADLSPLK 326
Db 317 LKIFKPEAAFLPENMGIAADLSPLK 342

RESULT 13
Q9TUD3 PRELIMINARY: PRT: 343 AA.
ID Q9TUD3
AC Q9TUD3;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyl transferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RL between an intronic Alu-Y element and red cell expression of ABH
RL antigens.";
RN Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080606; AAF14068.1; -.
DR InterPro; IPR002516; GP_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AB8C8 CRC64;

Query Match 73.78; Score 1396.5; DB 6; Length 343;
Best Local Similarity 79.4%; Pred. No. 1.1e-118;
Matches 259; Conservative 26; Mismatches 30; Indels 11; Gaps 2;

Oy 1 LQORIVKLQPLSEKELPMTTQSSNTSPESMRDSEOHNGELRGMTINSIGRLGNQM 60
Db 28 VQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGRLGNQM 76
Oy 61 GEYATLFALARMNGRLAFIPASMHNALAFIRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLYALAKMNGRPAFIPAQMHSTLAFIRITLPVLHSATASRIPWQNYHLNDWMEEE 136
Oy 121 YRHIPGHVVRTGTPCSTWTFYHHLRPEILKEFTLHDHVREEAQAFLRGLRVNGSPSTFV 180
Db 137 YRHIPGEVVRTGTPCSTWTFYHHLRQELQELQFTLHDHVREEAQAFLRGLRVNGSRPGTFV 196
Oy 181 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENINAS 240
Db 192 GVHVRG DYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCARENIDTS 256
Oy 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLAGGDTIYLANYLTPDSPF 300
Db 257 HGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGGIWAAYLTGGDTIYLANYLTPDSPF 316
Oy 301 LKVKPEAAFLPEWVGIPADLSPLK 326
Db 317 LKIFKPEAAFLPENMGIAADLSPLK 342

RESULT 14
Q29505 PRELIMINARY: PRT: 347 AA.
ID Q29505
AC Q29505;
DT 01-FEB-1997 (TremBLrel. 02, Created)
DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
```

01-MAR-2002 (TREMELREL. 20, Last annotation update)
 Galactoside 2-L-fucosyltransferase 3 (EC 2.4.1.69) (Secretor blood
 group ALPHA-2-fucosyltransferase) (GDP-L-fucose-beta-D-galactoside
 2-ALPHA-L-fucosyltransferase) (ALPHA(1,2)FUT) (Fucosyltransferase 3).
 Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GASTROINTESTINAL TRACT;
 RX MEDLINE=96279281; PubMed=8663168;
 RA Hitoshi S., Kojima K., Kanazawa I., Tsuji S.;
 RT "Molecular cloning and expression of a third type of rabbit GDP-L-
 RT fucose-beta-D-galactoside 2-alpha-L-fucosyltransferase.";
 RL J. Biol. Chem. 271:16975-16981(1996).
 CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE
 CC FUC-ALPHA(1,2)GALBETA- CALLED THE H ANTIGEN WHICH IS AN
 CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED
 CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE
 CC SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
 CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + BETA-D-GALACTOSYL-R = GDP -
 CC ALPHA-L-FUCOSYL-1,2-BETA-D-GALACTOSYL-R.
 CC -!- PATHWAY: GLYCOSYLATION.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI.
 CC -!- TISSUE SPECIFICITY: ADULT GASTROINTESTINAL TRACT AND COLON.
 CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
 CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
 CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
 CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
 CC GLYCOSYLTRANSFERASES.
 CC EMBL; X91269; CAA62669.1; -.
 DR InterPro: IPR002516; GT.11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack.
 FT DOMAIN 1 5
 FT TRANSMEM 6 28
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT LUMENAL, CATALYTIC (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 192 192
 FT CARBOHYD 286 286
 FT CARBOHYD 312 312
 FT CAROHOYD 312 312
 FT SEQUENCE 347 AA; 39469 MW; 746F7007309862A5 CRC64;
 Query Match 73.4%; Score 1391.5; DB 6; Length 347;
 Best Local Similarity 79.8%; Pred. No. 3.2e-118;
 Matches 260; Conservative 25; Mismatches 34; Indels 7; Gaps 1;
 QY 1 LQORIVKLPQLSKEKELPMTQMSSNGNTSEPMRDRSEQHGNGELRGMTINSIGRLGNQM 60
 DB 28 LQORLVRIQPTWEELPAL-----TPAVTRPTQSQRAPSLGSMWTINAMRLGNQM 80
 QY 61 GYATLFLALARMNGRLAFIPASMHNALAFIRISLPVLSHTAKKIPQWNYHLNDWMEER 120
 DB 81 GYATLYALAKENGRPAXTPAQMHSTLAFIRISLPVLSHTASTASRPVQWNYHLNDWMEER 140
 QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQPSTFV 180
 DB 141 YRHIPAPVRLTGYPCSWTFYHHLRHEILREFTLHDHVREEAQAFRLGRVNGSRPSTFV 200
 QY 181 GYVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMWCNENINAS 240
 DB 201 GYVRRGDYVHVMPQVWKGVADRGYLEKALDMFRARYSSPVFVTSNGMWCNENIDAS 260
 QY 241 RGDVVFAGNIEGSPAKDFALLTQCNTHTMTGTGFIWAAYLAGGDTIYLANVTLPDSPF 300
 DB 261 RGDVVFAGNLESPAKDFALLTQVNTHTMTGTGFIWAAYLTGGDTIYLANVTLPDSPF 320
 QY 301 LKVFKEPAFLPEWVGIPADLSPLK 326
 DB 321 LKIFKEPAFLPEWVGINADLSPLK 346

RESULT 15
 O77487
 ID 077487 PRELIMINARY; PST; 343 AA.
 AC 077487;
 DT 01-NOV-1998 (TREMELREL. 08, Created)
 DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)
 DT 01-MAR-2002 (TREMELREL. 20, Last annotation update)
 DE Alpha(1,2) fucosyltransferase.
 GN FUT2.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
 RT "The old origin of a null allele sc428 of the human ABO-secretor type
 RT alpha(1,2) fucosyltransferase gene (FUT2).";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95181460; PubMed=7876235;
 RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
 RT "Sequence and expression of a candidate for the human secretor blood
 RT group alpha(1,2) fucosyltransferase gene (FUT2).";
 RL J. Biol. Chem. 270:4640-4649(1995).
 DR EMBL; AB015636; BAA31129.1; -.
 DR InterPro: IPR002516; GT.11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 343 AA; 38973 MW; 2BD4D2A9704E4A0C CRC64;
 Query Match 73.3%; Score 1389.5; DB 6; Length 343;
 Best Local Similarity 79.1%; Pred. No. 4.8e-118;
 Matches 258; Conservative 27; Mismatches 30; Indels 11; Gaps 2;
 QY 1 LQORIVKLPQLSKEKELPMTQMSSNGNTSEPMRDRSEQHGNGELRGMTINSIGRLGNQM 60
 DB 28 LQORLAKIQAM--WELP-----EQIPVLASTSKALGPSOLRGITWTINATGLGNQM 76
 QY 61 GYATLFLALARMNGRLAFIPASMHNALAFIRISLPVLSHTAKKIPQWNYHLNDWMEER 120
 DB 77 GYATLYALAKENGRPAXTPAQMHSTLAFIRITLPLVLSHTTASRIPIQWNYHLNDWMEER 136
 QY 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQPSTFV 180
 DB 137 YRHIPGEYVRLTGYPCSWTFYHHLRHEILQEFTLHDHVREEAQAFRLGRVNGSQPSTFV 196
 QY 181 GYVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMWCNENINAS 240
 DB 197 GYVRRGDYVHVMPKVKGVADRGYLEKALDMFRARYSSPIFVTSNGMWCNENIDAS 256
 QY 241 RGDVVFAGNIEGSPAKDFALLTQCNTHTMTGTGFIWAAYLAGGDTIYLANVTLPDSPF 300
 DB 257 HSDVVFAGNIEGSPAKDFALLTQCNTHTMTGTGFIWAAYLAGGDTIYLANVTLPDSPF 316
 QY 301 LKVFKEPAFLPEWVGIPADLSPLK 326
 DB 317 LKIFKEPAFLPEWTGIAADLSPLK 342

Search completed: May 27, 2003, 15:10:46
 Job time : 38.6369 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:20:58 ; Search time 3098.72 Seconds
(without alignments)
10791.256 Million cell updates/sec

Title: US-10-040-863-7
Perfect score: 1149
Sequence: 1 atgcccagcgcccgagttcc.....gaagagccttcgatggaa 1149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rnd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1149	100.0	1149	10	AF264005	AF264005 Rattus no
2	1133.8	98.7	1847	10	AF131238	AF131238 Rattus no
3	1133.8	98.7	2984	10	AB006138	AB006138 Rattus no
4	1121.2	97.6	150840	2	AC121211	AC121211 Rattus no
5	1068	93.0	1068	10	AF042743	AF042743 Rattus no
6	931.8	81.1	6762	10	AF214656	AF214656 Mus muscu
7	931.8	81.1	200146	2	AC073774	AC073774 Mus muscu
8	865.8	75.4	1044	10	AF064792	AF064792 Mus muscu
9	802	69.8	963	10	AB039114	AB039114 Mus muscu
10	802	69.8	963	10	AB039115	AB039115 Mus muscu
11	802	69.8	963	10	AB039116	AB039116 Mus muscu
12	802	69.8	963	10	AB039117	AB039117 Mus muscu
13	802	69.8	963	10	AB039118	AB039118 Mus muscu
14	802	69.8	963	10	AB039119	AB039119 Mus muscu
15	802	69.8	963	10	AB039120	AB039120 Mus muscu
16	802	69.8	963	10	AB039121	AB039121 Mus spici
17	800.4	69.7	963	10	AB039122	AB039122 Mus muscu
18	798.8	69.5	963	10	AB039123	AB039123 Mus muscu
19	669.6	58.3	1555	10	AF131239	AF131239 Rattus no
20	659.6	57.4	1107	10	AB039213	AB039213 Mus spici
21	658	57.3	1107	10	AB039209	AB039209 Mus muscu
22	654.8	57.0	1107	10	AB039212	AB039212 Mus muscu
23	654.8	57.0	1107	10	AF113532	AF113532 Mus muscu
24	654.8	57.0	1589	10	MMSEC1	Y03882 M.musculus
25	653.2	56.8	1107	10	AB039206	AB039206 Mus muscu
26	653.2	56.8	1107	10	AB039207	AB039207 Mus muscu
27	653.2	56.8	1107	10	AB039208	AB039208 Mus muscu
28	653.2	56.8	1107	10	AB039211	AB039211 Mus muscu
29	653.2	56.8	4705	10	AF214657	AF214657 Mus muscu
30	651.6	56.7	1107	10	AB039210	AB039210 Mus muscu
31	650	56.6	1107	10	AB039204	AB039204 Mus muscu
32	650	56.6	1107	10	AB039205	AB039205 Mus muscu
33	647.6	56.4	1317	4	OCA12FT	X91269 O.cuniculus
34	633.8	55.2	2315	4	AF136895	AF136895 Sus scrof
35	633.8	55.2	2922	4	SSU070881	U70881 Sus scrofa
36	631.8	55.0	1044	4	AF027304	AF027304 Sus scrof
37	630.2	54.8	1043	6	AR212339	AR212339 Sequence
38	630.2	54.8	1043	6	AX029027	AX029027 Sequence
39	628.6	54.7	1043	6	AX029101	AX029101 Sequence
40	625.6	54.4	5357	10	AF214658	AF214658 Mus muscu
41	617.6	53.8	1062	9	AB015635	AB015635 Gorilla g
42	616	53.6	1062	9	AF080606	AF080606 Gorilla g
43	614.2	53.5	119631	9	AC008888	AC008888 Homo sapi
44	613.8	53.4	1784	9	BC001899	BC001899 Homo sapi
45	613.8	53.4	2115	6	AR039079	AR039079 Sequence

ALIGNMENTS

RESULT 1
AF264005
LOCUS
Rattus norvegicus alpha 1-2 fucosyltransferase mRNA, complete cds.
DEFINITION
AF264005
VERSION
AF264005.1 GI:8099676
KEYWORDS
SOURCE
Rattus norvegicus.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1149)
AUTHORS
Sherwood,A.L., Stroud,M.R., Levery,S.B. and Holmes,E.H.
TITLE
An amino acid region at the N-terminus of rat hepatoma alpha1-->2

fucosyltransferase modulates enzyme activity and interaction with
 lipids: strong preference for glycosphingolipids containing
 terminal Galbeta1->3GalNAc-structures
 Biochemistry 40 (19), 5708-5719 (2001)
 21240119
 11341836
 2 (bases 1 to 1149)
 Sherwood, A.L. and Holmes, E.H.
 Direct Submission
 Submitted (05-MAY-2000) Molecular Medicine, Northwest Hospital,
 2203 Airport Way South, Seattle, WA 98134, USA

FEATURES
SOURCE

CDS

```

/coon_start=1
/product="alpha 1-2 fucosyltrans
/protein_id="AAF72200.1"
/db xref="GI:80999677"

```

/translation="MASAOVPSPLAHLPLFVFTVSTIHLQRIVKLOPLSEKPL
 LTIQSGNTSPEDRSEHGELGKMTINIGLNGNGEYALHFLANNRNG
 LFIIPASMNALPRLSLVPLSDTAKTIPWNLHNGDMERYKLFHPHVRGT
 VCSWYTHLRLPEILKTHDHYREQAFLAGLRVNSQPTFVGVVRGDXVH
 VPMNVKGVADRGYELKALDMFRARYSPFVFTVSGMANGQPTINASRGDVFXAG
 GIESRGAOFALLTOCMTIMTIGTGVHAYLAGGDTIYLANTLPLDPSFLKVPKE
 AALPEWJZTIPDLSPLKALPACPGRSHHILKAGVCTYVAGRAF

BASE COUNT	247 a	341 c	311 g	250 t
ORIGIN				

```
Query Match      100.0%; Score 1149; DB 10; Length 1149;
Best Local Similarity 100.0%; Pred. No. 7.8e-297;
Matches 1149; Conservative 0; Mismatches 0; Indels 0;
```

	QY	I	ATGGCAGGCCGAGGTTCCTTTCTCCTTCCCTTGCCCACTTCCCATCTTTGTCTTC	60
	Db	1	ATGGCAGGCCGAGGTTCCTTTCTCCTTCCCTTGCCCACTTCCCATCTTTGTCTTC	60
	QY	61	GTGACTTCACCATCATCCACCCTCCAGCAGCGAATAGTGAAGCTCCAACCCCTGTGCAG	120
	Db	61	GTGACTTCACCATCATCCACCCTCCAGCAGCGAATAGTGAAGCTCCAACCCCTGTGCAG	120
	QY	121	AAGGAATTACCGATGACGACTCAAATGTCTCGGAAACACAGAAAAGCCCAGAGAIGGGA	180
	Db	121	AAGGAATTACCGATGACGACTCAAATGTCTCGGAAACACAGAAAAGCCCAGAGAIGGGA	180
	QY	181	CGGACAGCAGCAGCATGGGAATGGAGAGCTGCGGGGCATGTTTCAGATCAATTCCAAT	240
	Db	181	CGGACAGCAGCAGCATGGGAATGGAGAGCTGCGGGGCATGTTTCAGATCAATTCCAAT	240
	QY	241	GGCCGGCTGGGGAACAGATGGGGGAATAGCCACACATCTTTGCATGGCCAGAGTAAC	300
	Db	241	GGCCGGCTGGGGAACAGATGGGGGAATAGCCACACATCTTTGCATGGCCAGAGTAAC	300
	QY	301	GGACGGCTTCGGTTTCATCCCGCATCCATSCACAACGCTTAGCGCCCATCTTCAGGATC	360
	Db	301	GGACGGCTTCGGTTTCATCCCGCATCCATSCACAACGCTTAGCGCCCATCTTCAGGATC	360
	QY	361	AGCTCCCGGTGTTTACACAGCAGCACGGCCAAAAAGATCCCATGGCAGAAATACCATCTC	420
	Db	361	AGCTCCCGGTGTTTACACAGCAGCACGGCCAAAAAGATCCCATGGCAGAAATACCATCTC	420
	QY	421	AACGACTGGATGGAGGAGGHTTACCGCCACATTCGGGACACTTTGTCGGCTTCACGGGA	480
	Db	421	AACGACTGGATGGAGGAGGHTTACCGCCACATTCGGGACACTTTGTCGGCTTCACGGGA	480
	QY	481	TACCGTGTCTCGGACCTTCTACACACACTGGCCOCAGAGATCCTGAAGAGHTTACC	540
	Db	481	TACCGTGTCTCGGACCTTCTACACACACTGGCCOCAGAGATCCTGAAGAGHTTACC	540

QY	541	CTGCATGACCACGTCGGGAGGAGGCCCTCTCGTGGTCTGCGGGTGAAATGGG	600
Db	541	CTGCATGACCACGTCGGGAGGAGGCCCTCTCGTGGTCTGCGGGTGAAATGGG	600
QY	601	AGCCAGCCGAGTACTTTGTGGGTGTCCTAGTCGCGCCGAGGGGACATATGTCATGTCATG	660
Db	601	AGCCAGCCGAGTACTTTGTGGGTGTCCTAGTCGCGCCGAGGGGACATATGTCATGTCATG	660
QY	661	CCTAAATGTGTGGAAGGCGTGGTGGCTGACCGGGGTACCTGGAAGAGGCCCTGGGATATG	720
Db	661	CCTAAATGTGTGGAAGGCGTGGTGGCTGACCGGGGTACCTGGAAGAGGCCCTGGGATATG	720
QY	721	TTCCGGGCGACGCTATTCATCTCCAGTCTTCGAGTCTTCGAGTCTTCGAGTCTTCGAGTCT	780
Db	721	TTCCGGGCGACGCTATTCATCTCCAGTCTTCGAGTCTTCGAGTCTTCGAGTCTTCGAGTCT	780
QY	781	CGGAGGAACATTAATGCTTCCGAGGAGAGTGGTTCGCGGGCAATGGTATTTGAGGGG	840
Db	781	CGGAGGAACATTAATGCTTCCGAGGAGAGTGGTTCGCGGGCAATGGTATTTGAGGGG	840
QY	841	TCGCCAGCCAAAGGACTTCGCGCTGCCTCACCGAGTGCACACACACACCATCATGACTATTGGG	900
Db	841	TCGCCAGCCAAAGGACTTCGCGCTGCCTCACCGAGTGCACACACACACCATCATGACTATTGGG	900
QY	901	ACCTTTGGGATTTGGGCTGCCTACCTGGCAGTGGTGATACCAATACATGACTATTGGG	960
Db	901	ACCTTTGGGATTTGGGCTGCCTACCTGGCAGTGGTGATACCAATACATGACTATTGGG	960
QY	961	ACCTTTCCGATTTCTCGGTTCCTCAAGTCTTTAAGCCAGAGGAGCCCTTCCTACCGGAA	1020
Db	961	ACCTTTCCGATTTCTCGGTTCCTCAAGTCTTTAAGCCAGAGGAGCCCTTCCTACCGGAA	1020
QY	1021	TGGGTGGGCATCCCTGCGCATCTGTCCTTAAAGGCATTTAACCCAGAGCCTGTCTCT	1080
Db	1021	TGGGTGGGCATCCCTGCGCATCTGTCCTTAAAGGCATTTAACCCAGAGCCTGTCTCT	1080
QY	1081	CGGTCCCACTTCCACCTCAAGGCAAGGAGTGCATCTTTACGTCGAGGAGGAGCCTTC	1140
Db	1081	CGGTCCCACTTCCACCTCAAGGCAAGGAGTGCATCTTTACGTCGAGGAGGAGCCTTC	1140
QY	1141	TGATGGGAA 1149	
Db	1141	TGATGGGAA 1149	
RESULT 2			
AF131238			
LOCUS	AF131238	1847 bp	DNA linear
DEFINITION	Rattus norvegicus alpha 1,2-fucosyltransferase B (FTB) gene, complete cds.		
ACCESSION	AF131238		
VERSION	AF131238.1	GI:4580672	
KEYWORDS			
SOURCE	Rattus norvegicus.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 1847) Bureau,V., Marionneau,S., Caillaud-Thomas,A., Le Moullac-Vaidye,B., Liehr,T. and Le Pendu,J.		
TITLE	Comparison of the three rat GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferases FTA, FTB and FTC		
JOURNAL	Eur. J. Biochem. 268 (4), 1006-1019 (2001).		
MEDLINE	21099374		
PUBMED	11179967		
REFERENCE	2 (bases 1 to 1847)		
AUTHORS	Bureau,V., Le Moullac-Vaidye,B., Liehr,T., Denis,M. and Le Pendu,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-FEB-1999) INSERM U419, Institut de Biologie, 9 Quai Monceau, Nantes 44035, France		
FEATURES	Location/Qualifiers		
source	1..1847		

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Charitable Contributions</i>	16. <i>Charitable Contributions</i>
17. <i>Volunteer Work</i>	17. <i>Volunteer Work</i>
18. <i>Substance Use</i>	18. <i>Substance Use</i>
19. <i>Mental Health</i>	19. <i>Mental Health</i>
20. <i>Life Satisfaction</i>	20. <i>Life Satisfaction</i>

Db	811	AGCCAGCCGAGTACTTTTGTGGGTGTCCATGTGCGCGAGGGGACTATGTGCATGTGCATG	870
QY	661	CCTAATGTGTGGAAGGCGTGGTGGCTGACCGCGGGTTACCTGGAAAAGGCCCTGGATATG	720
Db	871	CCTAATGTGTGGAAGGCGTGGTGGCTGACCGCGGGTTACCTGGAAAAGGCCCTGGATATG	930
QY	721	TTCCGGSCACGCTATTATCATCTCCAGTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTGC	780
Db	931	TTCCGGSCACGCTATTATCATCTCCAGTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTGC	990
QY	781	CGSGAGCAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGGCAATGTAATGAGGGG	840
Db	991	CGSGAGCAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGGCAATGTAATGAGGGG	1050
QY	841	TCGCCAGCAAGACATTCGCGCTGCTCACCCAGTGCAACACACACACCATCATGACTATTGGG	900
Db	1051	TCGCCAGCAAGACATTCGCGCTGCTCACCCAGTGCAACACACACCATCATGACTATTGGG	1110
QY	901	ACCTTTGGGATTTGGGCTGCCTACCTGGCAGGTGGTGATACCACTACTTAGCCAACCTAC	960
Db	1111	ACCTTTGGGATTTGGGCTGCCTACCTGGCAGGTGGTGATACCACTACTTAGCCAACCTAC	1170
QY	961	ACCTTTCCGGATCTCGGTTCCTCAAAAGTCTTAAAGCAGAGGAGCGCTTCTTACCGGAA	1020
Db	1171	ACCTTTCCGGATCTCGGTTCCTCAAAAGTCTTAAAGCAGAGGAGCGCTTCTTACCGGAA	1230
QY	1021	TGGTGGGATCCCTGCGGATCTGTCCTCCACTCTTAAGGCACTTAACACGAGCTGTCCCT	1080
Db	1231	TGGTGGGATCCCTGCGGATCTGTCCTCCACTCTTAAGGCACTTAACACGAGCTGTCCCT	1290
QY	1081	CGGTCCCACTTCCACTCAAGGCAAAAGGAGTCACTTGTTCAGTCGAGGAGGAGCTTTC	1140
Db	1290	CGGTCCCACTTCCACTCAAGGCAAAAGGAGTCACTTGTTCAGTCGAGGAGGAGCTTTC	1349
QY	1141	TGATGGGAA 1149	
Db	1350	TGATGGGAA 1358	
RESULT 3			
AB006138			
LOCUS	AB006138	2984 bp	mrna
DEFINITION	Rattus norvegicus FTB mRNA for alpha 1,2-fucosyltransferase, complete cds.		linear ROD 05-FEB-1999
ACCESSION	AB006138.1	GI:2317265	
KEYWORDS	FTB; alpha 1,2-fucosyltransferase.		
SOURCE	Rattus norvegicus colon cancer cell_line:RCN-9 cDNA to mRNA.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (sites)		
TITLE	Soejima,M., Wang,B., Koda,Y. and Kimura,H.		
JOURNAL	Two distinct rat GSP-L-fucose-3-b-D-galactoside		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 2984)		
JOURNAL	Koda,Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (04-AUG-1997) Yoshiro Koda, Kurume University, School of		
JOURNAL	Medicine, Department of Forensic Medicine; Asahimachi 67, Kurume,		
FEATURES	Fukuoka 830, Japan (E-mail:ykoda@med.kurume-u.ac.jp,		
source	Tel:0942-31-7554, Fax:0942-31-7700)		
gene	Location/Qualifiers		
	1..2984		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/chromosome="1"		
	/map="1q22-q31"		
	/cell_line="RCN-9"		
	/tissue_type="colon cancer"		
	1..2984		

CDS	/gene="FTB" 214..1278 /gene="PTB" /codon_start=1 /product="alpha 1,2-fucosyltransferase" /protein_id="BAA21742.1" /db_xref="GI:2317266" /translation="MASAQVFPFPLAHLFIIVFTVSTIIHQORIVKLOPLSKELP LTFQSSGNTSPERMRDSEOHGSELRGMTINSIGRLGNOMGEYATLFLARMNGR MAFIPASMHNAALFIFLISLPVLSHSDAKKIPWQNYHLNDMEERYRHIPGHFVFTG YPCSWTYHHRLEPILFEFLTHDVRREAQFLKRLVNGSOPSTFGVHVHVRGDYH VMPNWGVVADRYILEKALDMFRATRYSSPVVYTSNGMAWNCINASRDGVVFNAG GIESPKAFDALLTQCHNIMTICTFGIWAAYLAGGDTIYLANTYLPLDPSPLKVEKE AAFLPEWVGIPADLSPLKHK"	729 a 808 c 761 g 586 t
BASE COUNT		
ORIGIN		
Query Match	98.7%; Score 1133.8; DB 10; Length 2984;	
Best Local Similarity	99.7%; Pred. No. 9.5e-293;	
Matches 1146; Conservative	0; Mismatches 2; Indels 1; Gaps 1;	
QY	1 ATGCCAGCGCCAGGTCCTTCCTTCCTCTGCGCCCACTTCCTCAICTTTGICTTC 60	
DB		
QY	214 ATGCCAGCGCCAGGTCCTTCCTTCCTCTGCGCCCACTTCCTCAICTTTGICTTC 273	
DB		
QY	61 GTGACTTCCACCATCATCCACTCCAGCAGCGGAATAGTGAAGCTCCCAACCCCTGTCCAGAG 120	
DB		
QY	274 GTGACTTCCACCATCATCCACTCCAGCAGCGGAATAGTGAAGCTCCCAACCCCTGTCCAGAG 333	
DB		
QY	121 AAGGAATTTACCGATGACGACATCAATGTCTCTGGGAACACAGAAAGCCAGAGATGGGA 180	
DB		
QY	334 AAGGAATTTACCGATGACGACATCAATGTCTCTGGGAACACAGAAAGCCAGAGATGGGA 393	
DB		
QY	181 CGGACACGCGAGCAGCATGGGAATGGAGAGCTGGGGGCATGTTCAACATCAATTCCTATT 240	
DB		
QY	394 CGGNCACGAGCAGCATGGGAATGGAGAGCTGGGGGCATGTTCAACATCAATTCCTATT 453	
DB		
QY	241 GCGCGCTGGGGAACACAGATGGGGAATACGCCACACTTTTGGCACTGGCCAGGATGAAC 300	
DB		
QY	454 GCGCGCTGGGGAACACAGATGGGGAATACGCCACACTTTTGGCACTGGCCAGGATGAAC 513	
DB		
QY	301 GGACGGCTTGGTTCATCCCGCGATCCATGCACAAAGCTCTAGCGCCCATCTTCAGGATC 360	
DB		
QY	514 GGACGGCTTGGTTCATCCCGCGATCCATGCACAAAGCTCTAGCGCCCATCTTCAGGATC 573	
DB		
QY	361 AGCCTCCGGTGTACACAGCGACACGCGCCAAAAGATCCCATGGCAGAAATACCATCTC 420	
DB		
QY	574 AGCCTCCGGTGTACACAGCGACACGCGCCAAAAGATCCCATGGCAGAAATACCATCTC 633	
DB		
QY	421 AAGCACTGGATGGAGAGCGTTACGGCCACATTCGGGGACACTTTGTGGCTTTCACGGGA 480	
DB		
QY	634 AACCACTGGATGGAGAGCGTTACGGCCACATTCGGGGACACTTTGTGGCTTTCACGGGA 693	
DB		
QY	481 TACCGTCTCTCTGACCTTCTACCAACACTGGCGCCAGAGATCCCTGAAGAGTTTACC 540	
DB		
QY	694 TACCGTCTCTCTGACCTTCTACCAACACTGGCGCCAGAGATCCCTGAAGAGTTTACC 753	
DB		
QY	541 CTGCATGACCACTGCGGGAGAGGSCCCAGCGCTTCCTCGGTGTCTCGGGGTGAATGGG 600	
DB		
QY	754 CTGCATGACCACTGCGGGAGAGGSCCCAGCGCTTCCTCGGTGTCTCGGGGTGAATGGG 813	
DB		
QY	601 AGCCAGCCGAGTACTTTTGTGGTGTCTCATGTGCGGCCAGGGACTATGTGCATGTCAATG 660	
DB		
QY	814 AGCCAGCCGAGTACTTTTGTGGTGTCTCATGTGCGGCCAGGGACTATGTGCATGTCAATG 873	
DB		
QY	661 CCTAATGTGGAAGGGGTGTGCTCACCGGGTTACCTGGAAAAGGCCCTTGGATATG 720	
DB		
QY	874 CCTAATGTGGAAGGGGTGTGCTCACCGGGTTACCTGGAAAAGGCCCTTGGATATG 933	
DB		
QY	721 TTCCGGGACACCTATTCATCTCCAGTCTTCGTGGTGTACAAAGCAGGATGTCCTGTGTC 780	
DB		
QY	934 TTCCGGGACACCTATTCATCTCCAGTCTTCGTGGTGTACAAAGCAGGATGTCCTGTGTC 993	
DB		
QY	781 CGGAGAACATTAATGCTTCCCGAGGAGACGTGGTGTTCGGGGGCAATGCTATTGAGGG 840	
DB		
QY	994 CGGAGAACATTAATGCTTCCCGAGGAGACGTGGTGTTCGGGGGCAATGCTATTGAGGG 1053	
DB		
QY	841 TCGCAGCAAGGACTTCCGCTGCTACCCAGTGCAACCAACACCATCATGACTATTGGG 900	
DB		
QY	1054 TCGCAGCAAGGACTTCCGCTGCTACCCAGTGCAACCAACACCATCATGACTATTGGG 1113	
DB		
QY	901 ACCTTTGGGATTTGGGCTGCTTACCTGCGAGTGGTGATACCATCTACTTACCAACTAC 960	
DB		
QY	1114 ACCTTTGGGATTTGGGCTGCTTACCTGCGAGTGGTGATACCATCTACTTACCAACTAC 1173	
DB		
QY	961 ACCTTTCCGGATTTCCCGCTTCCCTCAAAAGCTTTTAAGCCAGAGGAGCGCTTCTACCCGAA 1020	
DB		
QY	1174 ACCTTTCCGGATTTCCCGCTTCCCTCAAAAGCTTTTAAGCCAGAGGAGCGCTTCTACCCGAA 1233	
DB		
QY	1021 TGGTGGGCATCCCGGCGGATCTGTCCCACTCTCTTAAGGCAATTAACACGAGCTTGTCTT 1080	
DB		
QY	1234 TGGTGGGCATCCCGGCGGATCTGTCCCACTCTCTTAAGGCAATTAACACGAGCTTGTCTT 1292	
DB		
QY	1081 CGGTCCCACTTCCCACTCAAGGCAAGGAGTCACTTGTAGTGCAGGAGAGCCCTTC 1140	
DB		
QY	1293 CGGTCCCACTTCCCACTCAAGGCAAGGAGTCACTTGTAGTGCAGGAGAGCCCTTC 1352	
DB		
QY	1141 TGATGGGAA 1149	
DB		
QY	1353 TGATGGGAA 1361	
DB		
RESULT 4		
AC121211		
LOCUS	Rattus norvegicus clone CH230-309A14, *** SEQUENCING IN PROGRESS	
DEFINITION	***, 63 unordered pieces.	
ACCESSION	AC121211.2 GI:21909463	
VERSION	HTG: HTGS_PHASE1.	
KEYWORDS	Rattus norvegicus.	
SOURCE	Rattus norvegicus	
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 150840)	
AUTHORS	Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbarta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Davila,M.L., Davis,C., Coyle,M.D., Dathorne,S.R., David,R., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisseged,H., Lozados,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oraguyue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Qullies,M., Ren,Y.,	

Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
 Sodergren, E., Sonaika, J., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 150840)
 Worley, K.C.
 Direct Submission
 Submitted (16-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 150840)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 19, 2002 this sequence version replaced gi:20806241.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYKK
 Center clone name: CH230-309A14
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 106837 bases at least Q40
 Consensus quality: 113151 bases at least Q30
 Consensus quality: 116417 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 63 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1027: contig of 1027 bp in length
 * 1028 1127: gap of unknown length
 * 1128 2276: contig of 1149 bp in length
 * 2277 2376: gap of unknown length
 * 2377 3792: contig of 1416 bp in length
 * 3793 3892: gap of unknown length
 * 3893 4898: contig of 1006 bp in length
 * 4899 4998: gap of unknown length
 * 4999 6224: contig of 1226 bp in length
 * 6225 6325: gap of unknown length
 * 6326 7677: contig of 1353 bp in length
 * 7678 7777: gap of unknown length
 * 7778 8779: contig of 1002 bp in length
 * 8780 8879: gap of unknown length
 * 8880 10342: contig of 1463 bp in length
 * 10343 10442: gap of unknown length
 * 10443 11594: contig of 1152 bp in length
 * 11595 11694: gap of unknown length
 * 11695 13897: contig of 2203 bp in length
 * 13898 13997: gap of unknown length
 * 13998 15482: contig of 1485 bp in length
 * 15483 15582: gap of unknown length
 * 15583 15582: gap of unknown length
 * 15583 17010: contig of 1428 bp in length
 * 17011 17110: gap of unknown length
 * 17111 18427: contig of 1317 bp in length
 * 18428 18527: gap of unknown length
 * 18528 19811: contig of 1284 bp in length
 * 19812 19911: gap of unknown length
 * 19912 21647: contig of 1736 bp in length
 * 21648 21747: gap of unknown length
 * 21748 22753: contig of 1006 bp in length
 * 22754 22854: gap of unknown length
 * 22855 24607: contig of 1753 bp in length
 * 24608 24707: gap of unknown length
 * 24708 26168: contig of 1461 bp in length
 * 26169 26267: gap of unknown length
 * 26268 27443: contig of 1176 bp in length
 * 27444 27543: gap of unknown length
 * 27544 28740: contig of 1196 bp in length
 * 28741 28839: gap of unknown length
 * 28840 30991: contig of 2152 bp in length
 * 30992 32113: gap of unknown length
 * 32114 32213: contig of 1022 bp in length
 * 32214 33372: gap of unknown length
 * 33373 33472: contig of 1159 bp in length
 * 33473 33472: gap of unknown length
 * 33474 35105: contig of 1633 bp in length
 * 35106 35205: gap of unknown length
 * 35206 36378: contig of 1173 bp in length
 * 36379 36478: gap of unknown length
 * 36479 38090: contig of 1612 bp in length
 * 38091 38190: gap of unknown length
 * 38191 39491: contig of 1301 bp in length
 * 39492 41132: gap of unknown length
 * 41133 41232: contig of 1541 bp in length
 * 41233 42973: gap of unknown length
 * 42974 43073: contig of 1741 bp in length
 * 43074 44652: gap of unknown length
 * 44653 44752: gap of unknown length
 * 44753 45870: contig of 1118 bp in length
 * 45871 45970: gap of unknown length
 * 45971 48400: contig of 2430 bp in length
 * 48401 48500: gap of unknown length
 * 48501 50342: contig of 1842 bp in length
 * 50343 50442: gap of unknown length
 * 50443 52339: contig of 1897 bp in length
 * 52340 52439: gap of unknown length
 * 52440 53860: contig of 1421 bp in length
 * 53861 53960: gap of unknown length
 * 53961 56447: contig of 2487 bp in length
 * 56448 56547: gap of unknown length
 * 56548 57889: contig of 1342 bp in length
 * 57890 57989: gap of unknown length
 * 57990 59213: contig of 1224 bp in length
 * 59214 59313: gap of unknown length
 * 59314 61585: contig of 2272 bp in length
 * 61586 61685: gap of unknown length
 * 61686 63789: contig of 2104 bp in length
 * 63790 63889: gap of unknown length
 * 63890 65508: contig of 1619 bp in length
 * 65509 67142: gap of unknown length
 * 67143 67242: contig of 1534 bp in length
 * 67243 69536: contig of 2294 bp in length
 * 69537 69636: gap of unknown length
 * 69637 72776: contig of 3140 bp in length
 * 72777 72876: gap of unknown length
 * 72877 74143: contig of 1267 bp in length
 * 74144 74243: gap of unknown length
 * 74244 77710: contig of 3467 bp in length
 * 77711 77810: gap of unknown length
 * 77811 80194: contig of 2384 bp in length
 * 80195 80294: gap of unknown length
 * 80295 82739: contig of 2445 bp in length

RESULT 6	AF214656	6762 bp	DNA	linear	ROD 26-JUN-2001
LOCUS	AF214656				
DEFINITION	Mus musculus alpha(1,2)fucosyltransferase FUT2 (Fut2) gene, complete cds.				
ACCESSION	AF214656				
VERSION	AF214656.1	GI:7288504			
KEYWORDS					
SOURCE	Mus musculus.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 6762)				
AUTHORS	Domino,S.E., Zhang,L. and Lowe,J.B.				
TITLE	Molecular cloning, genomic mapping, and expression of two secretor blood group alpha(1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract				
JOURNAL	J. Biol. Chem. 276 (26), 23748-23756 (2001)				
MEDLINE	21316545				
PUBMED	11234119				
REFERENCE	2 (bases 1 to 6762)				
AUTHORS	Domino,S.E. and Lowe,J.B.				
TITLE	Mus musculus alpha(1,2)fucosyltransferase FUT2 (Sec2) genomic sequence				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 6762)				
AUTHORS	Domino,S.E. and Lowe,J.B.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med. Ctr. Dr., Ann Arbor, MI 48109-0650, USA				
FEATURES	Location/Qualifiers				
source	1..6762				
	/organism="Mus musculus"				
	/strain="129/Ola"				
	/db_xref="taxon:10090"				
	/chromosome="7"				
	/map="23.2; between Klkl and Ldh3"				
	/cell_line="E14"				
	/cell_type="ES"				
gene	<2210..>3253				
	/gene="Fut2"				
	/note="Sec2"				
mRNA	<2210..>3253				
	/gene="Fut2"				
	/product="alpha(1,2)fucosyltransferase FUT2"				
CDS	2210..3253				
	/gene="Fut2"				
	/EC_number="2.4.1.69"				
	/note="fucosyltransferase; similar to the Mus musculus strain ICR Sec2 sequence deposited at GenBank Accession Number AF064792; similar to the Homo sapiens secretor blood group alpha(1,2)fucosyltransferase FUT2 sequences deposited at GenBank Accession Numbers D87942 and U17894"				
	/codon_start=1				
	/product="alpha(1,2)fucosyltransferase FUT2"				
	/protein_id="AAF45146.1"				
	/db_xref="GI:7288505"				
	/translation="MASAQQPFSPLAHLFFVTSTIHLQORIVKLTSLSEKQLQAVMSNAPRFLRTDQMSAKLQGLFTINSIGRLNQMGEXATLFLARMNGLAFIPISYHMLALPISLPLVHSDTARIPWNYHLNOMDERYHILPGQVYFRTGYSQSWTFYHLRLEPILEKFTLHDSVREEAQIKGLRVNGSQSPFTVGVHVRGDIYVHVPKVKWFKGVADRGYLDKARLQYSSPVVYTSMAWCRNINTSLGDVVVFAGNGLEGSFKDPALITQCNNHTTNTIGTFCINAAYLAGGDTIYLIANYLTLPDSPFLKIFKPAAPLPEW MGIPADLSPLLKH"				
	4943..4948				
polya_signal	6333..6338				
polya_signal	1738 a 1682 c 1647 g 1691 t 4'others				
BASE COUNT					
ORIGIN					
Query Match	81.1%	Score 931.8	DB 10	Length 6762	

*	113115	117991: contig of 4877 bp in length	
*	117992	118091: gap of unknown length	
*	118092	146417: contig of 28326 bp in length	
*	146418	146517: gap of unknown length	
*	146518	154588: contig of 8071 bp in length	
*	154589	154688: gap of unknown length	
*	154689	158642: contig of 3954 bp in length	
*	158643	158742: gap of unknown length	
*	158743	160577: contig of 1835 bp in length	
*	160578	160677: gap of unknown length	
*	160678	196882: contig of 36205 bp in length	
*	196883	196982: gap of unknown length	
*	196983	200146: contig of 3164 bp in length.	
FEATURES			
Location/Qualifiers			
source	1..200146		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="RP23-36F17"		
	/clone_lib="RPCI mouse BAC library 23"		
BASE COUNT	47482 a	52711 c	50595 g 48057 t 1301 others
ORIGIN			
Query Match	81.1%; Score 931.8; DB 2; Length 200146;		
Best Local Similarity	89.7%; Pred. No. 1.5e-238;		
Matches 1031; Conservative	0; Mismatches 97; Indels 21; Gaps 2;		
QY	1	ATGCCACGCCAGGTCCTTCTCCTTCTCTGCCCCACTTCCTCATCTTTGTCCTTC	60
Db	14017	ATGGCGAGTGCCCAAGTACCTTCTCCTTCTCTGCCCCACTTCCTCATCTTTGTCCTTT	13958
QY	61	GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCCAACCCCTGTCAGAG	120
Db	13957	GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCCAACCCCTGTCAGAG	13898
QY	121	AAGGAATACCGATGACCACTCAAAATGCTCTCGGGAACACAGAAAGCCCGAGATGCGGA	180
Db	13897	AAGGAATATAC---AGGCGGTTCAAATGTCTCCACCAACGGCGCAAGACAGACAI----	13845
QY	181	CGGCACAGCAGCAGCATGGGAATGGAGAGCTGGGGGCATGTTACAGATCAATTCGATT	240
Db	13844	-----GCAGCAGATGCCAAGCTGCAGGCGCATATTACAGATCAATTCGATC	13799
QY	241	GGCGGCTGGGGAACACAGATGGCGGAATACGCCACTCTTTTGCACCTGCCCAGGATGAAC	300
Db	13798	GGCGGCTGGGGAACACAGATGGCGGAATATGCTACATGTTTGCACCTGCCCAGGATGAAC	13739
QY	301	GGACGGCTGGGTCATCCCGGATCCATGACACACGCTGAGCGCCCATCTTCAGGATC	360
Db	13738	GGTCGGCTGGGTCATCCCTGAAATCCATGACACACGCTCTAGCGCCCATCTTCAGGATC	13679
QY	361	AGCCTCCCGGTGTACACAGCGACACGCCCAAAAGATCCCATGGCAGAAATACCATCTC	420
Db	13678	AGTCTCCCGGTGTACACAGCGACACAGCCAGAGGATCCCGTGGCAGAAATACCATCTC	13619
QY	421	AACGACTGGATGGAGAGCGTTTACCGGCACATTCGGGACACTTTGTGCGCTTCACGGGA	480
Db	13618	AACGACTGGATGGAGAGCGTTTACCGGCACATTCGGGCGCATATGTCGTTTCACGGGA	13559
QY	481	TACCGGTGCTCTGTGACCTTCTACACACCTGGCCAGAGATCCCTGAAGAGGATTCACC	540
Db	13558	TACCGGTGCTCTGTGACCTTCTACACACCTGGCCAGAGATCCCTGAAGAGGATTCACC	13499
QY	541	CTGCATGACCAAGTCGGGGAGAGGCCAGCGCTTCCTGCTGCTGCGGGTGAATGGG	600
Db	13498	CTGCAGGACCATGTGCGTGAAGAGGCCAGCGTTCTTCTGCGTGGCGGTGAATGGG	13439
QY	601	AGCCAGCGAGTACTTTTGTGGGTGTCATGTGCGCGAGGGGACTATGTGCATGTGATG	660
Db	13438	AGCCAGCGGAGTACTTTTGTGGGTGTCATGTGCGCGAGGGGACTATGTGCATGTGATG	13379
QY	661	CCTAATGTGGAAGGCGGTGGTGGCTGACCGGGGTTACCTGGGAAAGGCCCTGGATG	720
Db	13378	CCCAAGGTGGAAGGCGGTGGTGGCTGACCGGGGTTACCTGGGAAAGGCCCTGGACAGG	13319

Qy	721	TTCCGGGCAACGCTA-TTCATCTCCAGTCTTCGTGGTTACAGCAACGGTATGCCCTGGTGC	780
Db	13318	TTCCGGGCAACGCTA-TTCATCTCCAGTCTTCGTGGTTACAGCAACGGTATGCCCTGGTGC	13259
Qy	781	CGGAGAAACATTAAATGCTTCCGAGGAGACGCTGGTGTTCGCGGGCAATGTTATGAGGGG	840
Db	13258	CGGAGAACATCAACACCTCCCTAGAGACGCTGGTGTTCGCGGGCAATGTTATGAGGGG	13199
Qy	841	TGCCAGCAAGGAGTCTCGCGCTGCTCCACCTGCTCAACCCAGTGCACACCATCATGATGGG	900
Db	13198	TCACCAGCAAGGAGTCTCGCGCTGCTCAACCCAGTGCACACCATCATGATGATGGG	13139
Qy	901	ACCTTTGGGATTTGGCTGCTTACCTGCGAGGTGTGATACCATCTACTTAGCCAACTAC	960
Db	13138	ACCTTTGGGATTTGGCTGCTTACCTGCGAGGTGTGATACCATCTACTTAGCCAACTAC	13079
Qy	961	ACCTTCGGGATTCCTCCCTCAAGTCTTAAAGCCAGAGGAGCGCTTCCTACCCGAA	1020
Db	13078	ACCTTCGGGATTCCTCCCTCAAGTCTTAAAGCCAGAGGAGCGCTTCCTACCCGAG	13019
Qy	1021	TGGGTGGCATCCCTGCGGATCTGTCCCATCTTAAAGCATTAACACCGCCCTGCT	1080
Db	13018	TGGATGGCATCCCTGCGGAGCTGTCCCATCTTAAAGCATTAACACCGCCCTGCT	12959
Qy	1081	CGGTCCCATCTCCACCTCAAGGCAAGAGAGTCACTTCTTACGTCGCGAGGAGCGCTTC	1140
Db	12958	TGGTCCCATCTCCACCTCAAGGCAAGAGTCACTTCTTACGTCGCGAGGAGCGCTTC	12899
Qy	1141	TGATGGGAA 1149	
Db	12898	TACTGGGAA 12890	
RESULT 8			
AF064792			
LOCUS			
DEFINITION Mus musculus GDP-L-fucose:beta-D-galactoside linear ROD 24-OCT-2000			
2-alpha-L-fucosyltransferase mRNA, complete cds.			
ACCESSION AF064792			
VERSION AF064792.1 GI:3142705			
KEYWORDS			
SOURCE Mus musculus.			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE 1 (bases 1 to 1044)			
AUTHORS Lin,B., Hayashi,Y., Saito,M., Sakakibara,Y., Yanagisawa,M. and Iwamori,M.			
TITLE GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II, and not MFUT-I or -III, is induced in a restricted region of the digestive tract of germ-free mice by host-microbe interactions and cycloheximide			
JOURNAL Biochim. Biophys. Acta 1487 (2-3), 275-285 (2000)			
MEDLINE 20471982			
PUBMED 11018479			
REFERENCE 2 (bases 1 to 1044)			
AUTHORS Lin,B., Hayashi,Y., Saito,M., Sakakibara,Y., Yanagisawa,M. and Iwamori,M.			
TITLE Direct Submission			
JOURNAL Submitted (12-MAY-1998) Chemistry, Faculty of Science & Technology, Kinki University, Kowakae, Higashi-Osaka 577-8502, Japan			
FEATURES			
Location/Qualifiers			
Source			
1..1044			
/organism="Mus musculus"			
/strain="ICR"			
/db_xref="taxon:10090"			
/tissue_type="gastrointestinal tract"			
1..1044			
/EC_number="2.4.1.69"			
/note="sec2; Se"			
/codon_start=1			
/product="GDP-L-fucose:beta-D-galactoside"			
CDS			


```

2-alpha-1-fucosyltransferase"
/db_xref="GI:3142706"
/translation="MASAOPFSEPLAHFLIFVFVSTIIHLQOQIVKLOTLSKEKIQ
AVQMSNNAARTDMQSAKIQGLFTINSIGRLNMGMEYATFLALANMGRFLAFIPES
MHLALPFRISLPLVHSDTARIPWNYHLNDWMEYRHPHPOQYVFTGYPCSWTF
YVHLRPELKEFTLHDHREAEAFGLRLVNGSQPSFVGVHVRGDYVHVMPKVMK
GVVADRGYLEKALDRARYSSPVVTSNGMAWCNENITSLGDVVFAGNLEGSPA
KDFALITQCNTHTMTICTFGIWAAYLAGGDTIVLANTLPDSPFLKIFKFAANFLPEW
MGIPADLSPLKH"
BASE COUNT      229 a 318 c 271 g 226 t
ORIGIN

Query Match      75.4%; Score 865.8; DB 10; Length 1044;
Best Local Similarity 90.3%; Pred. No. 6.1e-221;
Matches 956; Conservative 0; Mismatches 82; Indels 21; Gaps 2;

QY 1 ATGCCAGAGCCGACAGTTCCTTCTCTTCCCTGTGCCCCACTTCCTCATCTTTGTCTTC 60
DB 1 ATGGCGAGTCCCCAGGTACCTTTCTCTTCTCTTCCCTGTGCCCCACTTCCTCATCTTTGTCTTC 60
QY 61 GTGACTTCCACCATCATCCACCTCCACAGAGATAGTGAAGCTCCACACCCCTGTCCAG 120
DB 61 GTGACTTCCACCATCATCCACCTCCACAGAGATAGTGAAGCTCCACACCCCTGTCCAG 120
QY 121 AAGGAATACCGATGAGACTCAATATGCTCGGGAAACACAGAAAGCCGACAGATCCGA 180
DB 121 AAGGAATACCGATGAGACTCAATATGCTCGGGAAACACAGAAAGCCGACAGATCCGA 180
QY 181 CGGACAGCAGCAGCAGATGGATGAGAGCTCGGGGGGATGTTTCAGATCAATTCATT 240
DB 174 -----GCAGAGATGGCAAGCTCAGGGCATATTCAGATCAATTCATT 219
QY 241 GGCGGCTGGGAGACAGATGGGGAATAGCCACACTTTGACATGGCCAGGATGAAC 300
DB 220 GGCGGCTGGGAGACAGATGGGGAATAGCTATGTTGACATGGCCAGGATGAAC 279
QY 301 GGAGCGCTTCGGTTCATTCCTCCGCAATCCATGCACAAAGCTTAGCGCCATCTTCAGGATC 360
DB 280 GGTGCGCTTCGGTTCATTCCTCCGCAATCCATGCACAAAGCTTAGCGCCATCTTCAGGATC 339
QY 361 AGCTCCCGTGTATACACGCGACAGCGCCAAAGATCCCATGGCAGATTCACATCTC 420
DB 340 AGTCTCCCGTGTATACACGCGACAGCGCAAGGAACCGGTGGCAGATTCACATCTC 399
QY 421 AACGACTGATGAGGAGCGCTTACCGCCACATCCGCGACACTTTGCGCTTCACGGA 480
DB 400 AACGACTGATGAGGAGCGCTTACCGCCACATCCGCGGAGATGTCGCTTCACGGA 459
QY 481 TACCGTGTCTTCGGACCTTCATACGACAGCTCGGCCGACAGATCTGGAAGAGATCAAC 540
DB 460 TACCGTGTCTTCGGACCTTCATACGACAGCTCGGCCGACAGATCTGGAAGAGATCAAC 519
QY 541 CTGCATGACCACTGCGGAGGAGCGCCAGCGCTTCTCGTGTGCTGTCGCGGTGATGGG 600
DB 520 CTGCATGACCACTGCGGAGGAGCGCCAGCGCTTCTCGTGTGCTGTCGCGGTGATGGG 579
QY 601 AGCCAGCGAGTACTTTTGTGGTGTCCATGTGCGCGGAGGAGACTATGTCATGTGATG 660
DB 580 AGCCAGCGAGTACTTTTGTGGTGTCCATGTGCGCGGAGGAGACTATGTCATGTGATG 639
QY 661 CTTATGTTGGAGAGCGGTGTTGCGTACCGGGGTTACCTGAAAGGCCCTGGATG 720
DB 640 CTTATGTTGGAGAGCGGTGTTGCGTACCGGGGTTACCTGAAAGGCCCTGGATG 699
QY 721 TTCGGGACGCTTATCATCTCCAGTCTTCGTTGTTACAGACAGCGATGGCCTGGTGC 780
DB 700 TTCGGGACGCTTATCATCTCCAGTCTTCGTTGTTACAGACAGCGATGGCCTGGTGC 759
QY 781 CGGAGAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGCGCAATGTTATGAGGG 840
DB 760 CGGAGAACATTAATGCTTCCCGAGGAGAGCTGGTGTTCGCGGCGCAATGTTATGAGGG 819

```

```

QY 841 TCGCCAGCCAGGACITTCGGCTCTCACCAGTGCACACACACATCATCATATTGGG 900
DB 820 TCACAGCCAGGACITTCGGCTCTCACCAGTGCACACACACATCATCATATTGGG 879
QY 901 ACCTTTGGGATTTGGGCTGCTTACTTGCAGGTGGTGATACCATCTACTTACGCAACTAC 960
DB 880 ACCTTTGGGATTTGGGCTGCTTACTTGCAGGTGGTGATACCATCTACTTACGCAACTAC 939
QY 961 ACCCTTCCGATTTCCGTTCTCTCAAAAGTCTTTAAGCAGAGGAGCGCTTCTTACCGGAA 1020
DB 940 ACCCTTCCGATTTCCGTTCTCTCAAAAGTCTTTAAGCAGAGGAGCGCTTCTTCCCGAG 999
QY 1021 TGGTGGGATCCCTCGGATCTCTCCCACTCTCTTAAG 1059
DB 1000 TGGTGGGATCCCTCGGATCTCTCCCACTCTCTTAAG 1038

```

```

RESULT 9
AB039114
LOCUS      963 bp      DNA      linear      ROD 02-OCT-2001
DEFINITION Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
            2-alpha-1-fucosyltransferase, partial cds, strain:C57BL/10SnJ.
ACCESSION AB039114
VERSION    AB039114.1 GI:15822973
KEYWORDS   FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
SOURCE     Mus musculus (sub_species:domesticus, strain:C57BL/10SnJ) DNA.
ORGANISM   Mus musculus
            Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
            Liu,Y., Kitano,T., Koide,T., Shiomi,T., Moriaki,K. and
            Saitou,N.
            Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
            of Five Mus musculus subspecies
            Unpublished
            2 (bases 1 to 963)
            Liu,Y. and Saitou,N.
            Direct Submission
            Submitted (29-FEB-2000) Nariya Saitou, National Institute of
            Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,
            Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
            URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
            Fax:81-559-81-6789)
            Location/Qualifiers
                1..963
                /organism="Mus musculus"
                /strain="C57BL/10SnJ"
                /sub_species="domesticus"
                /db_xref="taxon:10090"
                /chromosome="7"
                /note="sequence used for primer design based on
                Acc#AF064792"
                1..963
                /gene="FUT2"
                <1..>963
                /gene="FUT2"
                /EC_number="2.4.1.63"
                /note="Sec2; Se"
                /codon_start=1
                /product="GDP-L-fucose:beta-D-galactoside
                2-alpha-1-fucosyltransferase"
                /protein_id="BAB68638.1"
                /db_xref="GI:15822974"
                /translation="FLIFVFTSTIIHQORIVKLTLSKEKIQAVQMSNNAARTDM
                QQSAKIQGLFTINSIGRLNMGMEYATFLALANMGRFLAFIPESMHLALPFRISLPLVHSDTARIPWNYHLNDWMEYRHPHPOQYVFTGYPCSWTFYVHLRPELKEFTL
                HDHREAEAFGLRLVNGSQPSFVGVHVRGDYVHVMPKVMKGVVADRGYLEKALDRARYSSPVVTSNGMAWCNENITSLGDVVFAGNLEGSPA
                KDFALITQCNTHTMTICTFGIWAAYLAGGDTIVLANTLPDSPFLKIFKFAANFLPEWM"
                TCTCTCAAAAGTCTTTAAGCAGAGGAGCGCTTCTTCCCGAG

```

```

BASE COUNT      214 a 285 c 257 g 207 t
ORIGIN

Query Match      69.8%; Score 802; DB 10; Length 963;

```


		CCACAGATGTTACCCCTGTCCGGGCATA	197
QY	223	TTCACGATCAATTCCAATTGGCGCGCTGGGGAACCAAGATGGGGAAATACGCCACACTCTT	282
Dd	160	TTCACGATCAATTCCAATCGCGCGCTGGGACCAGATGGCGAATATGCTACATTGTT	219

||||| 219


```

Db      460 ATCTTGAAGAGTTCACCTGACGACCATGTGCGTSGAGAGGCCAGGCTTCTCTCGT 519
QY      583 GGTCTCGGGTGAATGGGACGACGAGTACTTTTGTGGTGTCCATGTGCGCCAGGG 642
Db      520 GGCTCGGGTGAATGGGACGACGAGTACTTTTGTGGTGTCCATGTGCGCCAGGG 579
QY      643 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 702
Db      580 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTA 639
QY      703 GAAAGGCCCTGGATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 762
Db      640 GAAAGGCCCTGGAGAGTTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 699
QY      763 AACGATATGCTGTGTCGGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 822
Db      700 AACGATATGCTGTGTCGGGAGAACATTAATGCTCCGAGGAGACGTGTGTCGG 759
QY      823 GCATATGTCATGTCATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 882
Db      760 GCAATATGTCATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 819
QY      883 ACCATCATGATGTCATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 942
Db      820 ACCATCATGATGTCATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 879
QY      943 APTCTACTTAGCCAACTACACCTTCCGGATTCCTCGGTCCTCAAAGTCTTTAAGCCAGAG 1002
Db      880 APTCTACTTAGCCAACTACACCTTCCGGATTCCTCGGTCCTCAAAGTCTTTAAGCCAGAG 939
QY      1003 GCAGCTTCTTACCCGATGGGTG 1026
Db      940 GCAGCTTCTTACCCGATGGGTG 963

RESULT 12
LOCUS   AB039117
DEFINITION Mus musculus FUN2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase, partial cds, strain:CAST/Ei.
ACCESSION AB039117
VERSION   1
KEYWORDS  FUN2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
SOURCE   Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS  Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
TITLE    Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 963)
AUTHORS  Liu,Y. and Saitou,N.
DIRECT SUBMISSION
TITLE    Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES             Location/Qualifiers
     source           1..963
                     /organism="Mus musculus"
                     /strain="CAST/Ei"
     sub_species      "castaneus"
     db_xref          "taxon:10090"
     chromosome       "7"
     note             "sequence used for primer design based on
Acc#AF064792"
     1..963
     /gene="FUN2"

```

CDS

```

<1..>963
/gene="FUN2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB6841.1"
/db_xref="GI:15822980"
/translation="FLIFVFVSTIIHQQRIVKIQTLSEKELQAVOMSPNAARDTM
QSAKLOQITINSIGRLNQMEYATLALARMNGRLAFIPESMHNALAFIRISLP
VLHSDTARRIPWQNLHNDMEERYRHIPQVVRFTYPCSWTFYHHLPEILKETL
HDBVREAOAFLRLVNGSQPSTFVGVHVRGDDYVHVKVGVYVADRGYLEKALD
RRARYSSVFVFTSNGMAKRENINTSLGDVVFAGNGIEGSPAKDFALTGCTNHTIM
TIGTGTWAAIYLAGDTIYLANITLPSFLKIFKPAAPLPEWM"
BASE COUNT 214 a . 285 c 207 g 207 t
ORIGIN
Query Match 69.8%; Score 802; DB 10; Length 963;
Best Local Similarity 90.2%; Pred. No. 7.6e-204;
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;
QY 43 TTCTCATCTTTGCTCTGCTGACTTCCACCATCATCCACTCCAGAGGGAATAGTAGAG 102
Db 1 TTCTCATCTTTGCTCTTGTGACTTCCACCATCATCCACTCCAGAGGGAATAGTAGAG 60
QY 103 CTCCACCCCTCTCAGAGAGGAATACCGATGACGACTCAATGTCCTCGGGAGACACA 162
Db 61 CTCCACCCCTCTCAGAGAGGAATATAC---AGCGGTTCAAATGTCTCTCAACACGGG 117
QY 163 GAAAGCCAGAGATGCGACGGGACAGCGAGATGGGAATGGAGATGGGGGCATG 222
Db 118 GCAAGAACAGACAT-----GCAGCAGATGCCAAGCTGCAGGGCATA 159
QY 223 TTCAGATCAATTCATGCGCGGTGGGGAACACAGATGGGGGAATACGCCACACTTTT 282
Db 160 TTCAGATCAATTCATGCGCGGTGGGGAACACAGATGGGGGAATATGCTACATGTTT 219
QY 283 GCACGTGGCAGATGACGGAGCGGTGGGTTCATCCCGCGCATCCATGACACACCTCTA 342
Db 220 GCACGTGGCAGATGACGGGTGGGTTCATCCCTGGAATGCTGATGACACACCTCTA 279
QY 343 GGGCCATCTTCAGGATCAGCTCCCGGTGTGTACACAGCAGCAGCAGCAAAAGATCCCA 402
Db 280 GGGCCATCTTCAGGATCAGCTCCCGGTGTGTACACAGCAGCAGCAGCAGAGATCCG 339
QY 403 TGGCAGATTAACCATCTCAACGACTGGATGGAGAGCGTTACCGCACATTCGGGAGAC 462
Db 340 TGGCAGATTAACCATCTCAACGACTGGATGGAGAGCGTTACCGCACATTCGGGAGAG 399
QY 463 TTTGTGCGCTTCACGGATACCGGTCTCTGACCTTCTACACACCTGCGGCCAGAG 522
Db 400 TATGTGCGCTTCACGGATACCGGTCTCTGACCTTCTACACACCTGCGGCCAGAG 459
QY 523 ATCTGAAGAGTTCACCTGCAATGACCACTGCGGGAGGAGGCCAGCGCTTCCTCGGT 582
Db 460 ATCTGAAGAGTTCACCTGCAATGACCACTGCGGTGAGGAGGCCAGCGCTTCCTCGGT 519
QY 583 GGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTGGTGTCCATGTGCGCGAGGG 642
Db 520 GGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTGGTGTCCATGTGCGCGAGGG 579
QY 643 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTG 702
Db 580 GACTATGTCATGTCATGCTAATGTGTGGAAGGCGTGTGCTGACCGGGTTACCTA 639
QY 703 GAAAGGCCCTGGATGTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 762
Db 640 GAAAGGCCCTGGAGAGTTCGGGACGACCTATTATCTCCAGTCTTCCTGTTTACAGC 699
QY 763 AACGATATGCTGTGTCGGGAGAACATTAATGCTTCCGAGGAGAGACGTGTGTCGG 822
Db 700 AACGATATGCTGTGTCGGGAGAACATTAATGCTTCCGAGGAGACGTGTGTCGG 759

```



```

RESULT 14
AB039119
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase, partial cds, strain:pgn2.
ACCESSION
AB039119.1
VERSION
GI:15822983
KEYWORDS
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
SOURCE
Mus musculus (sub_species:domesticus, strain:pgn2) DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
JOURNAL
Unpublished
2 (bases 1 to 963)
Liu,Y. and Saitou,N.
Direct Submission
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Ilii Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="pgn2"
/sub_species="domesticus"
/db_xref="taxon:10090"
/chromosomes="7"
/note="sequence used for primer design based on
acc#AF064792"
1..963
/ gene="FUT2"
/ CDS
<1..>963
/ gene="FUT2"
/ EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase"
/protein_id="BAB68643.1"
/db_xref="GI:15822984"
/translation="FLIFVFTSTIIHQRIKVLQTLSEKELQVQSSNARTDM
QOSKLGIFINSGRLGNQGEYATLFLARMNGLAIPESMHALAIFRISLP
VLSDTARRIPQWYHNDMEERYHIPGYVRFYPCSWIFVHHLRPEILKEFTL
HDHVEEAQAFRLGLVNGSQSPSTFVGVVRRGDYVHVMPKWKGVADRGLKALD
RFRARYSSPVFVTVSNGMWCNRENTSLGVDVFPAGNIEGSPAKDFALLTQCNTIM
TIGTFGIWAAYLAGSDIYLYNLTLPDSPELKIKPAAFLPEWM"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 69.8%; Score 802; DB 10; Length 963;
Best Local Similarity 90.2%; Pred. No. 7.6e-204;
Matches 888; Conservative 0; Mismatches 75; Indels 21; Gaps 2;
QY 43 TTCCTCATCTTTGCTTGGTACCTCCACCATCCACCATCCACCGACGCGAATAGTGAAG 102
Db |||||||
1 TTCCCTCATCTTTGCTTGGTACCTCCACCATCCACCATCCACCGACGCGAATAGTGAAG 60
QY 103 CTCACACCCCTGTCAGAGAAGAAATACCGACTGACGACTCAAAATGCTCTCGGGAACACA 162
Db |||||||
61 CTCCAACCCCTGTCAGAGAAGAAATAC---AGCGGTTCAAATGCTCTCACCACCAACCGC 117
QY 163 GAAGCCGACAGATGCGGACGGACGACGACGATGGGAATGGAGAGCTGGGGGGCATG 222
Db |||||
118 GCAAGACAGACAT-----GCAGCAGAGTGCACAGCTCAGGGCATA 159
QY 223 TTCAGTCATTTCCATGGCGCTGGGGACACAGATGGCGCAATAGCCACACTCTTT 282
|||||

```

```

160 TTCACGTCATCAATCCATCGCGCGCTGGGGAACAGATGGCGGAATATGCTACATTTT 219
283 GCATCGCCAGGATGAACGAGCGCTTTCGCTTCATCCCGCATCCATGCAACAGCTCTA 342
|||||
220 GCATCGCCAGGATGAACGAGCGCTTTCGCTTCATCCCTGAATCCATGACACAGCTCTA 279
343 GCSCCCATCTTCAGGATCAGCTCCCGGTGTTACAGAGCGACAGCGCCAAAGATGCCA 402
Db |||||||
280 GCSCCCATCTTCAGGATCAGCTCCCGGTGTTACAGAGCGACAGCGCCAAAGATGCCG 339
403 TGCAGAAATACCATCTCAACGACTGGATGAGGAGGTTACCGCCACATTCGGGACAC 462
Db |||||||
340 TGCAGAAATACCATCTCAACGACTGGATGAGGAGGTTACCGCCACATTCGGGACAG 399
463 TTTGTGCGCTTCACGGGATACCGCTGCTCTCTGACCTTCTACACACCTTCGCGCCAGAG 522
|||||
400 TATGTGCGTTTCACGGGATACCGCTGCTCTCTGACCTTCTACACACCTTCGCGCCAGAG 459
523 ATCCTGAAGAGTTCACCGCTGCTATGACAGCGTGGGAGGAGGAGGCGCCAGGCTTCTCGCGT 582
Db |||||||
460 ATCCTGAAGAGTTCACCGCTGCTATGACAGCGTGGGAGGAGGAGGCGCCAGGCTTCTCGCGT 519
583 GGTCTCGCGTGAATGGGAGCGAGCTACTTTGTTGGTCTCATCTGCGCCGAGCG 642
Db |||||||
520 GGCCTCGCGTGAATGGGAGCGAGCTACTTTGTTGGTCTCATCTGCGCCGAGG 579
643 GACTATGTCATGTCATGCTTAATGTGTGAAGGCGTGTGCTGACCGGGGTACCTG 702
Db |||||||
580 GACTATGTCATGTCATGCTTAATGTGTGAAGGCGTGTGCTGACCGGGGTACCTA 639
QY 703 GAAAGGCCCTGGATATGTTCCGGGACAGCTATTATCTCCAGTCTTCGTGGTTACAAG 762
Db |||||||
640 GAAAGGCCCTGGATATGTTCCGGGACAGCTATTATCTCCAGTCTTCGTGGTTACAAG 699
763 AACGTTATGCGCTGTGCGGGAGAACATTAATGCTTCCGAGGAGCTGTTGTCGCG 822
Db |||||||
700 AACGTTATGCGCTGTGCGGGAGAACATTAATGCTTCCGAGGAGCTGTTGTTGCG 759
QY 823 GCATATGTTGAGGGTTCGCCAGCAAGGACTTCGCGCTGCTACCCAGTGCACACAC 882
Db |||||||
760 GGCATATGTTGAGGGTTCGCCAGCAAGGACTTCGCGCTGCTACCCAGTGCACACAC 819
QY 883 ACCATCATGACTATGCGGACCTTGGGATTTGGGCTGCTACCTGCGAGGTTGGTAIAC 942
Db |||||||
820 ACCATCATGACTATGCGGACCTTGGGATTTGGGCTGCTACCTGCGAGGTTGGTAIAC 879
QY 943 ATCTACTAGCAACTACACCTTCCGATTCCTGCTCCTCAAGTCTTTAAGCCAGAG 1002
Db |||||||
880 ATCTACTAGCAACTACACCTTCCGATTCCTGCTCCTCAAGTCTTTAAGCCAGAG 939
1003 GCAGGCTTCCCTACCCGAATGGTG 1026
Db |||||||
940 GCAGGCTTCCCTACCCGAATGGTG 963
RESULT 15
AB039120
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase, partial cds, strain:SWN/Msf.
ACCESSION
AB039120.1
VERSION
GI:15822985
KEYWORDS
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
SOURCE
Mus musculus (sub_species:molossinus, strain:SWN/Msf) DNA.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Liu,Y., Kitano,T., Koide,T., Shirolishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
JOURNAL
Unpublished

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 09:17:27 ; Search time 298.004 Seconds
(without alignments)
8682.922 Million cell updates/sec

Title: US-10-040-863-7

Perfect score: 1149

Sequence: 1 atggccagcgcccgattcc.....gaagagccttcgatgggaa 1149

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1149	100.0	1149 21	AC67965
2	1149	100.0	1149 24	AA27207
3	1068	93.0	1068 21	AA67966
4	1068	93.0	1068 24	AA27208
5	630.2	54.8	1043 19	AA29003
6	630.2	54.8	1043 19	AA21639
7	613.8	53.4	2115 19	AAV58323
8	612.6	53.3	3088 24	ABL64735
9	612.6	53.3	3088 24	ABL66311
				Rat hepatoma H35 c
				Rat hepatoma H35 c
				Rat hepatoma H35 c
				Rat hepatoma H35 c
				Porcine secretor t
				Pig secretor encod
				Human Sec2 coding
				Lung cancer relate
				Lung cancer relate

10	612.6	53.3	3088	24	ABL70026	Pancreas cancer re
11	423.8	36.9	1098	19	AAV21640	Pig H transferase
12	423.8	36.9	1269	20	AA15872	Fragment of PH83x
13	423.8	36.9	1269	20	AA15872	Swine alpha-1,2-fu
14	423.8	36.9	1269	21	AA294417	Pig alpha-1,2-fuco
15	423.8	36.9	1269	21	AA294417	FUT1 cDNA. Sus sc
16	405	35.2	1155	16	AAQ98461	GDP-L-fucose-beta-
17	405	35.2	1155	16	AAQ98461	2-Alpha-fucosyltra
18	405	35.2	1155	16	AAQ98461	Human H-transferas
19	405	35.2	1174	17	AA12238	Human H-transferas
20	405	35.2	1174	17	AA12238	Mouse H2kb gene pr
21	405	35.2	2268	18	AA12238	Human alpha 1,2 fu
22	405	35.2	3373	18	AA12238	Chicken beta-actin
23	405	35.2	3791	18	AA12238	GDP-Puc-beta-D-gal
24	405	35.2	8174	12	AA12238	Human alpha(1,2)-f
25	405	35.2	8174	18	AA12238	DNA encoding a gly
26	403.4	35.1	8174	15	AA12238	X. laevis alpha-1,
27	304.8	26.5	1071	20	AA12238	Human secreted pro
28	141.4	12.3	473	21	AA12238	Rabbit alpha1-2fuc
29	72.8	6.3	100	21	AA12238	Rabbit alpha1-2fuc
30	71.2	6.2	100	21	AA12238	Rabbit alpha1-2fuc
31	71.2	6.2	100	21	AA12238	Rabbit alpha1-2fuc
32	69.6	6.1	100	21	AA12238	Human alpha1-2fuc
33	63.2	5.5	100	21	AA12238	Human alpha1-2fuc
34	63.2	5.5	100	21	AA12238	Human alpha1-2fuc
35	55.8	4.9	100	21	AA12238	Human alpha1-2fuc
36	49.8	4.3	79	21	AA12238	Rabbit alpha1-2fuc
37	44	3.8	495	22	AA12238	Human cDNA clone (
38	42.8	3.7	4069	22	AA12238	Human cDNA sequenc
39	42.8	3.7	5236	22	AA12238	Human polynucleoti
40	42.8	3.7	5236	22	AA12238	Human polynucleoti
41	40.4	3.5	2297	22	AA12238	Human secreted pro
42	40.4	3.5	2612	22	AA12238	Human secreted pro
43	40.4	3.5	2953	22	AA12238	DNA encoding novel
44	40.4	3.5	3044	22	AA12238	Human cDNA encodin
45	40.4	3.5	3044	24	ABQ65608	Human polynuc.leoti

ALIGNMENTS

RESULT 1

AC67965
ID AAC67965 standard; cDNA; 1149 BP.

XX AAC67965;

AC AC

19-FEB-2001 (first entry)

Rat hepatoma H35 cell alpha1-2fucosyltransferase cDNA.

Rat; alpha1-2fucosyltransferase; cytotstatic; neuroprotective;

nootropic; gene therapy; Fucalalpha-2Galbeta1-3GalNAc; immunotherapy;

immunosuppression; cancer; neurological disease;

small cell lung carcinoma; ss.

Rattus norvegicus.

XX WO200064464-A1.

PD 02-NOV-2000.

XX 23-APR-1999; 99WO-US07384.

XX 23-APR-1999; 99WO-US07384.

(PACI-) PACIFIC NORTHWEST CANCER FOUND.

Holmes EH, Sherwood AL;

WPI: 2000-687262/57.

P-PSDB; AAB36104.

or immunotherapeutic for cancer and neurological diseases -

Claim 11; Fig 3A; 91pp; English.

The present sequence is given in a specification relating to a rat ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein or its cellular fraction is useful for synthesis of a molecule comprising Fucal-2Gal-beta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein or a free oligosaccharide comprising Fucal-2Gal-beta1-3GalNAc. The method involves contacting alpha1-2fucosyltransferase with GDP-fucose and a molecule or glycolipid, glycoprotein, glycolipoprotein or oligosaccharide having a terminal Gal-beta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM1 by contacting the protein with GDP-fucose and ganglioside GM1. The obtained glycoproteins, glycolipoproteins, glycolipids and oligosaccharides are useful as nutritional compositions and fucosyl-GM1 is useful for inducing an immunotherapeutic or immunosuppressive action against cancer, neurological disease or small cell lung carcinoma.

Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 93.0%; Score 1068; DB 21; Length 1068;
Best Local Similarity 100.0%; Pred. No. 7,3e-288;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 82 CTCGACGAGCGAATAGTGAAGCTCCACCCCTGTCCAGAGAAGGAATTACCGATGACGACT 141
Db 1 CTCGACGAGCGAATAGTGAAGCTCCACCCCTGTCCAGAGAAGGAATTACCGATGACGACT 60

QY 142 CAATGTCTCTCGGGAACACACAGAACGCCAGATGGACGGGACAGCAGCATGGG 201
Db 51 CAATGTCTCTCGGGAACACACAGAACGCCAGATGGACGGGACAGCAGCATGGG 120

QY 202 AATGAGAGCGTCGGGGGATCTTCCAGCATCAATTCATTCGCGCGTGGGGAACACAGATG 261
Db 121 AATGAGAGCGTCGGGGGATCTTCCAGCATCAATTCATTCGCGCGTGGGGAACACAGATG 180

QY 262 GGCGAATACGCCACACCTCTTTCGACTGCCAGGATGAACGACGCGCTGCTTCATCCCC 321
Db 181 GGCGAATACGCCACACCTCTTTCGACTGCCAGGATGAACGACGCGCTGCTTCATCCCC 240

QY 322 GCATCCATGACACAGCTCTAGCGCCCATCTCTCAGGATCAGCCCTCCGCGTGTACACAGC 381
Db 241 GCATCCATGACACAGCTCTAGCGCCCATCTCTCAGGATCAGCCCTCCGCGTGTACACAGC 300

QY 382 GACACGCCCAAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 441
Db 301 GACACGCCCAAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 360

QY 442 TACGCCACATTCGGGACACCTTTGTGGCTTCACGGATACCGCTGCTCTGGACCTTC 501
Db 361 TACGCCACATTCGGGACACCTTTGTGGCTTCACGGATACCGCTGCTCTGGACCTTC 420

QY 502 TACCACACCTGCGCCACAGATCTCTAAGGAGTTCACCTGCAATGACACCGTCCGGAG 561
Db 421 TACCACACCTGCGCCACAGATCTCTAAGGAGTTCACCTGCAATGACACCGTCCGGAG 480

QY 562 GAGCCCGCAGGCTTCCTCGCTGCTCGGCTGAATGGAGCCAGCCGAGTACTTTGTG 621
Db 481 GAGCCCGCAGGCTTCCTCGCTGCTCGGCTGAATGGAGCCAGCCGAGTACTTTGTG 540

QY 622 GGTGTCCATGTGCCCGGAGGACTATGTGATGTCATGTCCTAATGTGTGAAGGCGGTG 681
Db 541 GGTGTCCATGTGCCCGGAGGAGTATGTGATGTCATGTCCTAATGTGTGAAGGCGGTG 600

QY 682 GTGGCTGACGGGTTTACCTGGAAGGCCCTGGATATGTCGGGACCGCTATTCATCT 741
Db 601 GTGGCTGACGGGTTTACCTGGAAGGCCCTGGATATGTCGGGACCGCTATTCATCT 660

QY 742 CCACTCTTCGTGTTACAAGCAACGGTATGCCCTGTGTCGGGGAACATTAATGCTTCC 801
Db 661 CCACTCTTCGTGTTACAAGCAACGGTATGCCCTGTGTCGGGGAACATTAATGCTTCC 720

```

```

QY 802 CGAGGAGACGTGGTGTTCGGGGCAATGGTATTGAGGGTCCGACGACCAAGCACTTCGGC 861
Db 721 CGAGGAGACGTGGTGTTCGGGGCAATGGTATTGAGGGTCCGACGACCAAGCACTTCGGC 780

QY 862 CTGCTTCACCCAGTGCACACCAACCATCATGACTATTGGACCTTTGGGATTGGGCTGCC 921
Db 781 CTGCTTCACCCAGTGCACACCAACCATCATGACTATTGGACCTTTGGGATTGGGCTGCC 840

QY 922 TACCTGGAGGTGGTATACCATCTACTTAGCAACTACACCTTCGGGATTCTCCGTTTC 981
Db 841 TACCTGGAGGTGGTATACCATCTACTTAGCAACTACACCTTCGGGATTCTCCGTTTC 900

QY 982 CTCGAAGTCTTTAAGCCAGAGACGCTTCCTTACCCGAATGGTGGGATCCCTCCGAT 1041
Db 901 CTCGAAGTCTTTAAGCCAGAGACGCTTCCTTACCCGAATGGTGGGATCCCTCCGAT 960

QY 1042 CTGTCCCGACCTCCTTAGGCAATTAACACACGCTTCTCGGTCCGATCCACCTCAAG 1101
Db 961 CTGTCCCGACCTCCTTAGGCAATTAACACACGCTTCTCGGTCCGATCCACCTCAAG 1020

QY 1102 GCARAAGAGTCACTTGTAGTCCGACGAGAGAGCCCTTCTGATGGAA 1149
Db 1021 GCARAAGAGTCACTTGTAGTCCGACGAGAGAGCCCTTCTGATGGAA 1068

RESULT 4
RAD27208
ID AAD27208 standard; DNA; 1068 BP.
XX
AC AAD27208;
XX
DT 09-APR-2002 (first entry)
XX
DE Rat hepatoma H35 cell alpha1-2FucT catalytic domain encoding DNA.
XX
KW Rat; alpha1-2fucosyltransferase; alpha1-2FucT; antisense therapy;
KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytosstatic; ganglioside; GM1;
KW cell transformation; catalytic domain; ds.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1..1062
FT /tag= a
FT /product= "Rat hepatoma H35 cell alpha1-2FucT catalytic
FT domain"
FT /note= "CDS does not include start codon"
FT /partial
FT 581..1068
FT /tag= b
FT /note= "Region which overlaps rat FTB"
XX
US6329170-B1.
XX
PN 11-DEC-2001.
XX
PD 23-APR-1999; 99US-0298886.
XX
PF 23-APR-1999; 99US-0298886.
XX
PR (NWHO-) NORTHWEST HOSPITAL.
XX
PI Holmes EH, Sherwood AL;
XX
XX WPI; 2002-121132/16.
XX DR P-PSDB; AAE16623.
XX
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
XX GM1-specific alpha1-2fucosyltransferase enzyme by recombinant
XX techniques and for detecting oncogenic transformation of test tissues -
XX

```

Claim 2; Fig 3; 41pp; English.

The invention relates to rat GM1-specific alpha1-2fucosyltransferase (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic acid is specific for a carbohydrate moiety found in ganglioside GM1, a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc) saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the preparative synthesis of fucosyl containing glycolipids, glycoproteins, glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1. Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which involves assaying for changes in expression of alpha1-2 Fuct. Since alpha1-2Fuct is activated in cell transformation, antisense sequences derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and antisense therapy. The present sequence is rat hepatoma H35 cell alpha1-2Fuct catalytic domain encoding DNA.

Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 93.0%; Score 1068; DB 24; Length 1068;
Best Local Similarity 100.0%; Pred. No. 7.3e-289;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CTCGAGCGCAATAGTGAAGCTCCAAACCCCTGTGTCAGAGAGGAATTACCGAIGACGACT 141
DB 1 CTCGAGCGCAATAGTGAAGCTCCAAACCCCTGTGTCAGAGAGGAATTACCGAIGACGACT 60
QY 142 CAAATGTCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGGAGCAGCATGGG 201
DB 61 CAAATGTCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGGAGCAGCATGGG 120
QY 202 AATGAGAGCTCGGGGCAATGTTACGATCAATTCATTTGGCGGCTGGGGAACAGATG 261
DB 121 AATGAGAGCTCGGGGCAATGTTACGATCAATTCATTTGGCGGCTGGGGAACAGATG 180
QY 262 GCGCAATAGGCACACTCTTTCACCTGGCAGGATGAACGAGCGGCTGCGTTATCCGCC 321
DB 181 GCGCAATAGGCACACTCTTTCACCTGGCAGGATGAACGAGCGGCTGCGTTATCCGCC 240
QY 322 GCATCCATGCACAGGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTACACAGC 381
DB 241 GCATCCATGCACAGGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTACACAGC 300
QY 382 GACAGGCCAAAGATCCCATGGCAGAAATACCATCTCAAGCACTGGATGAGGAGCGGT 441
DB 301 GACAGGCCAAAGATCCCATGGCAGAAATACCATCTCAAGCACTGGATGAGGAGCGGT 360
QY 442 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGATACCCCGTCTCTGGACCTTC 501
DB 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGATACCCCGTCTCTGGACCTTC 420
QY 502 TACCACCACTCGCGCCAGAGATCTGAAGGAGTTCAACCTGTCATGACACGCTGCGGGAG 561
DB 421 TACCACCACTCGCGCCAGAGATCTGAAGGAGTTCAACCTGTCATGACACGCTGCGGGAG 480
QY 562 GAGGCCCAAGCGCTTCTCGCTGGTCTGCGGGTGAATGGAGCCACCGAGTACTTTTGTG 621
DB 481 GAGGCCCAAGCGCTTCTCGCTGGTCTGCGGGTGAATGGAGCCACCGAGTACTTTTGTG 540
QY 622 GGTGTCCATGTGCGCGGAGGAGTATGTGATGTCATGCTTAAATGTGGAAGGCGGTG 681
DB 541 GGTGTCCATGTGCGCGGAGGAGTATGTGATGTCATGCTTAAATGTGGAAGGCGGTG 600
QY 682 GTGGCTGACCGGGTTACTTGAAAGGCCCTCGGATATGTTCCGGSCACGCTATTTCATCT 741
DB 601 GTGGCTGACCGGGTTACTTGAAAGGCCCTCGGATATGTTCCGGSCACGCTATTTCATCT 660
QY 742 CCAGTCTTCTGGTTTACAGCAACGGTATGGCCCTGGTCCCGGAGAACATTAATGCTTCC 801
DB 661 CCAGTCTTCTGGTTTACAGCAACGGTATGGCCCTGGTCCCGGAGAACATTAATGCTTCC 720
QY 802 CGAGGAGAGCTGGTGTTCGCGGGGCAATGGTATTGAGGGGTGCGCCAGCCAGGACTTCGCG 861

DB 721 CGAGGAGAGCTGGTGTTCGCGGGCAATGTAATGAGGGTTCGCCAGGAGACTTCGCG 780
QY 862 CTGCTACCCAGTGGCAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 921
DB 781 CTGCTACCCAGTGGCAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 840
QY 922 TACCTGGCAGGTGGTGATACCATCTACTTACGCAACTACACCTTCGCGATTCGCGTTC 981
DB 841 TACCTGGCAGGTGGTGATACCATCTACTTACGCAACTACACCTTCGCGATTCGCGTTC 900
QY 982 CTCAAAGTCTTTAAGCCAGAGCAGCCCTTCTACCCGAATGGGTGGGCATCCCTTGCCGAT 1041
DB 901 CTCAAAGTCTTTAAGCCAGAGCAGCCCTTCTACCCGAATGGGTGGGCATCCCTTGCCGAT 960
QY 1042 CTGTCCCACTCTTTAAGGCATTAACACCACTTAACACCACTGTCCCTCGGTCCCACTCAAG 1101
DB 961 CTGTCCCACTCTTTAAGGCATTAACACCACTGTCCCTCGGTCCCACTCAAG 1020
QY 1102 GCAAAGGAGTCACTTTGTTACGTTCGAGGAAGAGCCCTTCTGATGGAA 1149
DB 1021 GCAAAGGAGTCACTTTGTTACGTTCGAGGAAGAGCCCTTCTGATGGAA 1068
RESULT 5
AAV29003
ID AAV29003 standard; DNA; 1043 BP.
XX
AC AAV29003;
XX
DT 28-AUG-1998 (first entry)
XX
DE Porcine secretor transferase (FUT2) gene.
DE
KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;
KW transgenic animal; xenotransplantation; organ transplant; ss.
XX
OS Sus scrofa.
XX
Key Location/Qualifiers
CDS 9..1031
/*tag= a
XX
PN WO9807837-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-AU00540.
XX
PR 23-AUG-1996; 96AU-0001823.
XX
PA (AUST-) AUSTIN RES INST.
XX
PI McKenzie IPC, Sandrin MS;
XX
DR WPI; 1998-169148/15.
XX
P-PSDB; AAW37855.
XX
PT Nucleic acid encoding glycosyltransferase able to compete with
PT second such enzyme - particularly used to reduce expression of
PT unwanted carbohydrate epitope(s) on tissues intended for
PT transplantation
XX
PS Claim 6; Fig 1A-B; 40pp; English.
XX
CC This nucleotide sequence, designated clone pSe16.1, codes for
CC porcine secretor (Se) glycosyltransferase (see AAW37855), an enzyme
CC that has high affinity for type I and type III substrates. It was
CC isolated from a pig liver genomic library using full-length human
CC Se2 (FUT2) cDNA as probe. The nucleotide sequence shows about 36%
CC homology with human FUT1. Introduction of glycosyltransferase
CC nucleic acid, such as porcine Se nucleic acid, into a tissue
CC results in reduced expression of unwanted carbohydrate epitopes on

[illegible]

```

QY 1021 TGGGTGGGATCCCTGCGGATCTGTGCCACACCTTTAAG 1059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 987 TGGATTGGGATCGAGGACAGACCTGTCCCACTCCTTAAG 1025

RESULT 7
AAV58323
ID AAV58323 standard; DNA; 2115 BP.
XX AC AC
XX XX
XX XX
XX XX
XX XX
XX DE Human Sec2 coding sequence.
XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
KW FUT2; nonsecretor genotyping; ds.
XX OS Homo sapiens.
XX XX
XX Key Location/Qualifiers
XX CDS 64..1095
XX FT /*tag= a
XX FT
XX PN US5807732-A.
XX XX
XX XX 15-SEP-1998.
XX XX 28-FEB-1995; 95US-0395800.
XX PF
XX XX 28-FEB-1995; 95US-0395800.
XX PR
XX XX {GIOR/} GIORGI D.
XX PA {KELL/} KELLY R J.
XX PA {LENN/} LENNON G.
XX PA {LOWE/} LOWE J B.
XX PA {ROUQ/} ROQUIER S.
XX XX
XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
XX WPI; 1998-520137/44.
XX P-PSDB; AAW69332.
XX DR
XX XX
XX PT DNA encoding fucosyltransferase enzyme - useful for producing
XX recombinant enzyme and genotyping person as secretor or nonsecretor
XX XX
XX Claim 1; Column 45-50; 55pp; English.
XX XX
XX This sequence encodes the human Sec2 protein of the invention. The DNA
XX encodes a alpha(1,2) fucosyltransferase and is the Secretor
XX alpha(1,2)fucosyltransferase locus, that cross hybridises with the
XX H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
XX producing a recombinant human GDP-L-fucose:beta-D-galactoside
XX 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
XX individual as a secretor or nonsecretor as it is known that nonsecretors
XX homozygous for a mutant allele of the FUT2 gene that has a stop codon in
XX the position corresponding to amino acid 143.
XX SQ Sequence 2115 BP; 505 A; 606 C; 552 G; 452 T; 0 other:

```

[illegible]

Claim 1; SEQ ID 3072; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour.

Sequence 3088 BP; 723 A; 839 C; 798 G; 728 T; 0 other;

Query Match 53.3%; Score 612.6; DB 24; Length 3088;

Best Local Similarity 73.8%; Pred. No. 1.4e-160;

Matches 808; Conservative 0; Mismatches 254; Indels 33; Gaps 1;

Qy	9	CGCCAGGTCCTTCTCCTTCCCTTCCCTGCGCCACTTCATCTTGTGCTGACTTC	68
Db	120	CGTTACAGATGCTTTCCTTCCCTTCCCTGCGCCACTTCATCTTGTGCTGACTTC	179
Qy	69	CACCATCATCACCTCCAGAGCGCAATAGTGAAGCTCCAAACCCCTGTACAGAGGAAT	128
Db	180	CACATATATTTACAGCTTCAGAGCGCGTAGCGAAGATTCAAGCCATGTGGAGTTAC	239
Qy	129	ACCATGACGACTCAATGTCTCGGGAACACAGAGCCAGAGATGCGACGGACAG	188
Db	240	GCAGTACCACTGTAGCTCAATCAAGCAAGCTGCGGACCCA-----	283
Qy	189	CGAGCAGCATGGGAATGGAGCTGCGGGGATGTCACGATCAATTCATTTGCGCGGCT	248
Db	284	-----GCCAGCTCAGGGGATGCGGACGATCAATGCAATAGCGCGCT	326
Qy	249	GGGAAACAGATGGGGAATACGCCAATCTTTGCTGCTGCGGAGGATGAACGAGCGCT	308
Db	327	GGGAAACAGATGGGGAATACGCCAATCTTTGCTGCTGCGGAGGATGAACGAGCGCT	386
Qy	309	TGCGTTCATCCCGCATCCATGACAGCTCTAGCGCCATCTCAGGATCAGCCCTCC	368
Db	387	CGCCTTATCCCGCATCCATGACAGCTCTAGCGCCATCTCAGGATCAGCCCTCC	446
Qy	369	GGTGTACACAGCAGCGGCGCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTG	428
Db	447	GGTGTACACAGCAGCGGCGCAAGATCCCATGGCAGAAATACCATCTCAACGACTG	506
Qy	429	GATGGAGAGCGTTACGCCACATTCGCGGACATTTGTGCGCTTCAGGGATACCGCTG	488
Db	507	GATGGAGAGCGTTACGCCACATTCGCGGAGTACGTCGCGCTTCAGCGGCTACCGCTG	566
Qy	489	CTCCTGGACCTTCTACCAACCTTCGCCCAAGATCCTGAGGAGTTCACCTTCATGA	548
Db	567	CTCCTGGACCTTCTACCAACCTTCGCCCAAGATCCTGAGGAGTTCACCTTCATGA	626
Qy	549	CCAGTGGCGGAGGAGGCGCCAGGCTTCTCGTGGTGTGCGGGTGAATGGAGCGACCC	608
Db	627	CCAGTGGCGGAGGAGGCGCCAGGCTTCTCGTGGTGTGCGGGTGAATGGAGCGACCC	686
Qy	609	GAGTACTTTTGTGGGTGCTTCATGTCGCCGAGGAGTATGTGATGTCAATGCTTAATGT	668
Db	687	GGGCAACCTTTGTGGGTGCTTCATGTCGCCGAGGAGTATGTGATGTCAATGCTTAATGT	746
Qy	669	GTGGAGGCGGCTGTGGTGTACCGGGGTGTACTGGAAAGGCCCTGGATGTTCGCGGC	728
Db	747	GTGGAGGCGGCTGTGGTGTACCGGGGTGTACTGGAAAGGCCCTGGATGTTCGCGGC	806

Qy	729	ACGCTATTATCTCAGTCTTCGTTTACAGCAACGGTATGCTGCTGCCGGAGAA	788
Db	807	TCGTACAGCTCCCTCATCTTCGTTGCTCACCAGTAAATGGCATGGCTGGTGTCCGGAGAA	866
Qy	789	CATTAATGCTTCCGAGAGGACGAGTGTGTTTCGCGGGCAATGTAATGAGGGTCCGCCAGC	848
Db	867	CATTGACACCTTCCACGCTGATGTTGCTGGCGATGSCATTGAGGGCTACCTGC	926
Qy	849	CAAGGACTTCGGGTGCTCACCAGTGCACACACACCATCATGACTATGGACCTTTGG	908
Db	927	CAAGATTTTGTCTTACATCACAGTGTAAACACACCATCATGACTATGGACCTTTGG	986
Qy	909	GATTGGCTGCTCCTACCTGGCAGGTGGTATACCATCTTACTTAGCCAACTACACCTTCC	968
Db	987	GATCTGGGCGCATACCTCAGCGGGGAGACCACTTACTTGGCAATTACACCTTCC	1046
Qy	969	GGATTCTCGTTCCTCAAAATCTTTAAGCCAGAGCGCTTCTTACCCGAATGGTGGG	1028
Db	1047	CGACTCCCTTTCTCTCAAAATCTTTAAGCCAGAGCGCTTCTTACCCGAATGGTGGG	1106
Qy	1029	CATCCCTGCGATCTGTCCCACTCTTAAAGGCATTACACACGCTTCTCTCGTCCCA	1088
Db	1107	GATTGCGCAGACCTGTCCCTTACTCAAGCACTAATGCTGGCCATTCTTIGAGACT	1166
Qy	1089	CTTCCACCTCAAGGC	1103
Db	1167	TTTCTCTCTCTGC	1181
RESULT 9			
ABL66311	ID ABL66311 standard; DNA; 3088 BP.		
XX	AC	ABL66311;	
XX	DT	15-MAY-2002 (first entry)	
XX	DE	Lung cancer related gene sequence SEQ ID NO:4648.	
XX	KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
XX	KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;	
XX	KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;	
XX	OS	gene; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO200194629-A2.	
XX	PD	13-DEC-2001.	
XX	PF	30-MAY-2001; 2001WO-US10838.	
XX	PR	05-JUN-2000; 2000US-209473P.	
XX	PR	05-JUN-2000; 2000US-209531P.	
XX	PR	18-SEP-2000; 2000US-233133P.	
XX	PR	18-SEP-2000; 2000US-233617P.	
XX	PR	20-SEP-2000; 2000US-234009P.	
XX	PR	20-SEP-2000; 2000US-234034P.	
XX	PR	20-SEP-2000; 2000US-234052P.	
XX	PR	22-SEP-2000; 2000US-234509P.	
XX	PR	22-SEP-2000; 2000US-234567P.	
XX	PR	25-SEP-2000; 2000US-234923P.	
XX	PR	25-SEP-2000; 2000US-234924P.	
XX	PR	25-SEP-2000; 2000US-235077P.	
XX	PR	25-SEP-2000; 2000US-235082P.	
XX	PR	25-SEP-2000; 2000US-235134P.	
XX	PR	25-SEP-2000; 2000US-235280P.	
XX	PR	26-SEP-2000; 2000US-235637P.	
XX	PR	26-SEP-2000; 2000US-235638P.	
XX	PR	27-SEP-2000; 2000US-235711P.	
XX	PR	27-SEP-2000; 2000US-235720P.	
XX	PR	27-SEP-2000; 2000US-235840P.	

P-PSDB; AAY79302.
Improving weight gain in swine using swine genetically resistant
Escherichia coli and feeding swine high levels of plant based protein
Disclosure; Fig 1; 33pp; English.
This is the nucleotide sequence of the pig FUT1 gene encoding
alpha-1-2 fucosyltransferase (see AAY79302). A polymorphism at
position 307 of the coding region is associated with susceptibility
to F18 Escherichia coli colonization; pigs homozygous for adenine
at position 307 are resistant to colonization, while heterozygous
animals and animals homozygous for guanine at position 307 are
susceptible to colonization. A claimed method for improving weight
gain in pigs involves selecting animals that are genetically
resistant to E. coli colonization and feeding these animals high
levels of plant-based protein concentrate. A claimed method for
preventing F18 E. coli colonization in swine, especially swine that
are genetically susceptible to F18 E. coli colonization, involves
replacing some or all of the plant-based proteins in the diet with
animal-based proteins. The polymorphism in the FUT1 gene is also
useful for developing drugs to treat swine that have E. coli
associated disease. The polymorphism can be detected using
PCR-RFLP tests (see also AA294418-19).
Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other.
Query Match 36.9%; Score 423.8; DB 21; Length 1269;
Best Local Similarity 69.6%; Pred. No. 5e-108;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;
217 GCGATGTTACGATCAATTCATTCATGGCGGCTGGGGAACACAGATGGGGAATACGCCACA 276
243 GGGACCTGGACTATTACCGGATGGCGCGTTGGGAACAGATGGGACATATGCCACG 302
277 CTCCTTGGCACTGGCAGATGAACGAGCGGTGGTTCATCCCGCATCCATGCCACAAC 336
303 CTGCTGGCCCTGGGCGAGCTCAACGGCGCCAGGCGCTTCATCCAGCCTGCCATGACGCC 362
337 GCTCTAGCGCCCACTCAGATCAGCTCCCGGTGTACACGGGACAGCGCCAAAAG 396
363 GTCCGCGCCCGGTGTCGCGATCAGCTGCTGCTGCGCGCGGAGGTAGACAGGCAC 422
397 ATCCATGGCAGAAATTACATCTCAACACTGGATGGAGGAGCGTTACCGCACATTCGG 456
423 GCTCTGGCGGAGCTGGAGCTCAACGACTGGATGTCGGAGGATTATGCCACTTAAG 482
457 GGACACTTGTGGGCTTACGGGATACCGGTGCTGCTGGACCTTCTACACACACTGGCG 516
483 GAGCCCTGGCTGAAGCTACACGGCTTCCCTGCTGCGCGCGGAGGTAGACAGGCAC 542
517 CCAGAGATCCTGAAGGAGTTCACCTGCATGACACGCTGCGGAGGAGGCCAGGCCCTTC 576
543 GAGCATGCTCCAGCGAGTTTACCTTGCACGACCACTTCGCGAAGAGGCCCGGGGTA 602
577 TTGCTGCTGCTGCGGCTGAATGGGAGCGAGCGAGTACTTTGTGGGTGTCCAI 630
603 CTGAGTCAGTTCCTGCTACCCCGCACAGGGAGCCGCCACCACTTCGTGGGGTCCAC 662
631 GTGGCGGAGGGGACTATGTCATGTGTCATGCTTAATGTGTGAAGCGCGTGGTGTGAC 690
663 GTGGCGCGCGGGACTATCTGCTGATGATGCCAAGCGCTGGAAGGGGTGTGGGTGAC 722
691 CGGGTTCACCTGGAAAGGCCCTGATATGTTCGGGCACTTATTCATCTCCAGCTTTC 750
723 GCGCTTACCTCCAGAGGCTATGACGTGTTCGGGCGCCGATAGAGCCCGGCTTT 782
751 GTGTTTACAACAACGGTATGCTGCTGGCGGAGAACATTAATGCTTCCCGAGGAC 810
783 GTGGTCACACAGCGGATGGAGTGGTGGCGGAGAACATCGACACCTCCCGGGGAC 842
811 GTGGTGTTCGCGGGCAATGTTATGAGGGGTTCGCCAGCAAGGACTTCGCGCTGCTCACC 870

843 GTGATCTTTGTCGGCATGGCGGAGGCCGCCGCCAGGACATTTGCGCTGCTGGTS 902
871 CAGTGAACCAACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTCCCTACCTGGCA 930
903 CAGTGAACCAACACCATCATGACTATTGGGACCTTTGCGCTTCTGGGCCCTACCTGGCT 962
931 GGTGTGATACCATCTACTTAGCCACTACACCCCTCCGGATTCCTCCGTTCCCAAGTC 990
963 GGTGAGATACCATCTACTTAGCCACTACACCCCTCCGGATTCCTCCGTTCCGATGTC 1022
991 TTAAAGCCAGAGCGAGCCCTTCTACCCGAATGGTGGGATCCCTGCCGATCTGTCCTCA 1050
1023 TTAAAGCCAGAGCGAGCCCTTCTACCCGAATGGTGGGATTAATGACACTTGTCTCCA 1082
1051 CTCCTTAAG 1059
1083 CTCAGATG 1091
RESULT 15
AAI72831
ID AAI72831 standard; cDNA; 1269 BP.
AC AAI72831;
DT 22-JUL-2002 (first entry)
XX
DE FUT1 cDNA.
KW Gene; pig; swine; alpha (1,2) fucosyltransferase; FUT1; weight gain;
F18; Escherichia coli; ss.
XX
OS Sus scrofa.
FH Key Location/Qualifiers
FT CDS 9..1106
FT /tag= a
FT /product= "FUT1"
FT mutation 315
FT /tag= b
FT /phenotype= "Confers resistance to F18 E. coli"
XX US6355859-B1.
PN 12-MAR-2002.
PD XX
XX 18-SEP-1998; 980S-0151592.
PF XX
PR 20-MAY-1997; 970S-047181P.
XX
PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
PA (USDA) US SEC OF AGRIC.
XX
PI Bosworth B, Ridpath J, Wiseman B;
XX
DR WPI; 2002-391652/42.
DR P-PSDB; AAB47995.
XX
PT Feeding swine genetically susceptible to F18 Escherichia coli
PT colonization with a low plant protein based diet increases weight gain
PT and lowers intestinal disease associated with E. coli infection
XX
PS Claim 1; Column 13-18; 9pp; English.
XX
CC This sequence represents the swine alpha (1,2) fucosyltransferase
CC (FUT1) gene. A FUT1 gene in which there is a base other than adenine
CC at position 307, may be used for improving weight gain in swine that
CC are genetically susceptible to F18 Escherichia coli. The weight gain
CC may be activated by feeding a diet of at least 40% animal based
CC proteins. The feeding method is used to control F18 E. coli associated
CC intestinal disease in swine.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 05:50:26 ; Search time 62.7104 Seconds
(without alignments)
5819.028 Million cell updates/sec

Title: US-10-040-863-7
Perfect score: 1149
Sequence: 1 atggccagcgcccggttc.....gaagagcctctgatggaa 1149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCrUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1149	100.0	1149	4	US-09-298-886-7
2	1068	93.0	1068	4	US-09-298-886-9
3	630.2	54.8	1043	4	US-09-254-077A-5
4	613.8	53.4	2115	1	US-08-395-800A-7
5	516.6	45.0	1144	1	US-08-395-800A-1
6	423.8	36.9	1269	4	US-09-151-592-1
7	408.2	35.5	1155	1	US-08-208-889A-1
8	408.2	35.5	1155	2	US-08-433-271-1
9	405	35.2	999	5	PCT-US91-00899-10
10	405	35.2	1136	1	US-08-395-800A-9
11	405	35.2	1174	5	PCT-US95-07554-3
12	405	35.2	1199	1	US-08-395-800A-5
13	405	35.2	2268	4	US-08-675-773B-4
14	405	35.2	3373	1	US-08-273-411-2
15	405	35.2	3791	4	US-08-675-773B-3
16	405	35.2	8174	1	US-07-914-281-5
17	405	35.2	8174	1	US-08-393-246-5
18	405	35.2	8174	1	US-08-526-058A-5
19	405	35.2	8174	2	US-08-696-731-5
20	405	35.2	8174	4	US-09-042-531-5
21	405	35.2	8174	5	PCT-US91-00899-3
22	403.4	35.1	1155	1	US-08-434-151-1
23	403.4	35.1	1155	2	US-08-715-259-1
24	72.8	6.3	100	4	US-09-298-886-29
25	71.2	6.2	100	4	US-09-298-886-23
26	71.2	6.2	100	4	US-09-298-886-26
27	59.6	6.1	100	4	US-09-298-886-20

US-09-298-886-7

Sequence 7, Application US/09298886

Patent No. 6329170

GENERAL INFORMATION:

APPLICANT: ERIC H. Holmes et al.

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 8511-029

CURRENT APPLICATION NUMBER: US/09/298,886

CURRENT FILING DATE: 1999-04-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 1149

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1143)

US-09-298-886-7

ALIGNMENTS

RESULT 1

US-09-298-886-7

Sequence 7, Application US/09298886

Patent No. 6329170

GENERAL INFORMATION:

APPLICANT: ERIC H. Holmes et al.

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 8511-029

CURRENT APPLICATION NUMBER: US/09/298,886

CURRENT FILING DATE: 1999-04-26

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 1149

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1143)

US-09-298-886-7

Query Match 100.0%; Score 1149; DB 4; Length 1149;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCCAGCGCCAGGTTCCCTTCTCTTCTCTGCGCCACTTCTCATCTTTGTCTTC 60

Db 1 ATGGCCAGCGCCAGGTTCCCTTCTCTTCTCTGCGCCACTTCTCATCTTTGTCTTC 60

Qy 51 GTGACTTCCACCATCATCCACCTCCAGCAGCGAATAGTGAAGCTCAACCCCTGTTCAG 120

Db 51 GTGACTTCCACCATCATCCACCTCCAGCAGCGAATAGTGAAGCTCAACCCCTGTTCAG 120

Qy 121 AAGGAATACCGATGACGACTCAATGTCTCTCGGAAACACAGAACGCCAGAGATGGA 180

Db 121 AAGGAATACCGATGACGACTCAATGTCTCTCGGAAACACAGAACGCCAGAGATGGA 180

Qy 181 CGGACAGCGACACATGGGAATGGAGCTGCGGGCATGTTCCAGCATCAATTCATT 240

Db 181 CGGACAGCGACACATGGGAATGGAGCTGCGGGCATGTTCCAGCATCAATTCATT 240

Qy 241 GCGCGGCTGGGAAACAGATGGCGAATACGCCACACTCTTTGCACCTGCCAGGATGAC 300

Db 241 GCGCGGCTGGGAAACAGATGGCGAATACGCCACACTCTTTGCACCTGCCAGGATGAC 300

Qy 301 GGACGGCTGGCTTCCATCCCGCATCCATGACACACGCTTAGCGCCCATCTTCAGGATC 360

Db 301 GGACGGCTGGCTTCCATCCCGCATCCATGACACACGCTTAGCGCCCATCTTCAGGATC 360

Db 1083 CTCAGATG 1091

RESULT 7

US-08-208-889A-1
Sequence 1, Application US/0820889A
Patent No. 5750176
GENERAL INFORMATION:
APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,
APPLICANT: Moremen, K., Pierce, J.
TITLE OF INVENTION: Transgenic Production of Oligosaccharides and
TITLE OF INVENTION: Glycoconjugates
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Donald O. Nickey
ADDRESSEE: ROSS Products Division
ADDRESSEE: Abbott Laboratories
STREET: 525 Cleveland Avenue
CITY: Columbus
STATE: Ohio
COUNTRY: United States of America
ZIP: 43215
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS Version 6.21
SOFTWARE: WordPerfect Version 6.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208,889A
FILING DATE: 09-MAR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA: No. 5750176 applicable
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614) 624-7080
TELEFAX: (614) 624-3074
TELEX: No. 5750176e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Cloned cDNA representing the product of a
MOLECULE TYPE: human genomic DNA segment
DESCRIPTION: GDP-L-fucose- D-galactoside 2-alpha-fucosyl-transferase
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Entire amino acid sequence provided.
ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 19
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: DNA sequencing and restriction analysis
OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:
OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- D-galactoside 2-alpha-fucosyl-
OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO: 1.
OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.

Db 992 GCGGAGACACTGCTACCTGGCCAACTTCACTCCGACACTGAGTTCTCTGAAGATC 1051
Qy 991 TTTAAGCCAGGAGGAGCCCTTCTTACCGGATGGTGGGATCTGTCGCCATCTGTCGCCA 1050
Db 1052 TTTAAGCCAGGAGGAGCCCTTCTTACCGGATGGTGGGATCTGTCGCCATCTGTCGCCA 1111
Qy 1051 CTC 1053
Db 1112 CTC 1114

RESULT 9
PCT-US91-00899-10
; Sequence 10, Application PC/TUS9100899
; GENERAL INFORMATION:
; APPLICANT: Lowe, John B.
; TITLE OF INVENTION: Method and Products For the Synthesis of
; TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,
; TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned
; TITLE OF INVENTION: Genetic Sequences That Determine these Structures
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00899
; FILING DATE: 19910214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye Ph.D., Jean-Paul
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
; TELEPHONE: (703)521-5940
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Blood
; PCT-US91-00899-10

Query Match 35.2%; Score 405; DB 5; Length 999;
Best Local Similarity 68.4%; Pred. No. 2.1e-108;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

Qy 217 GGCATGTTACAGTCAATTCATTCATGGCGGTGGGACACAGATGGGGAATAGCCACA 276
Db 139 GGCACCTGGATGTTACCCCAATGGCGGTGTTGTAATCAGATGGGACATGCCCAG 198
Qy 277 CTCCTTGACCTGGCCAGGATGAACGAGCGGCTTGGTTTCATCCCGCATCCATGCACAC 336
Db 199 CTGCTGGCTTGGCCCACTACAGCGCGCGGCGCTTATCCCTGCTGCTGCTGCTGCTG 258
Qy 337 GCTCTAGCGCCCATGTTCTGAGGATAGCTCCCGGTGTTTACAGGACACGCGGCAAAAG 396

Db 259 GCCTTGCCCGCGGTATTCCTGCATCACCTGCCCTGCTGCCCCAGAGTGAGCAGCCGC 318
Qy 397 ATCCATGGCAGATTAACCATCTCAACGACTGGATGGAGGAGCGTTACCGCCACATTCG 456
Db 319 AGCCGTGGCGGAGCTGACGCTTCACGACTGGATGTCGGAGGAGTACGGGAGCTTGA 378
Qy 457 GGACACTTTTGGCGCTTCAGGGATACCCGCTGCTTCCCTGGACCTTCTTACCAACCTCC 516
Db 379 GATCCTTTCTGAAAGCTCTCTGGCTTCCCTGCTTCTTGGACTTCTTCCACCATCTCC 438
Qy 517 CCAGAGATCTCTGAAGGAGTTCACCCCTGCATGACACGTCGGGAGGAGGCCGCCCTTC 575
Db 439 GAACAGATCCGAGAGATTCACCCCTGCACGACACCTTCGGGAGAGGCGACAGTGTG 498
Qy 577 CTGCTGCTGTCGGGTG-----AATGGGAGCGAGCTACTTTTGTGGGTGTCCAT 630
Db 499 CTGGGTCAGCTCCCTTGGCGGACAGGGGACCGCCGCGACCTTGTGGGCTCCAC 558
Qy 631 GTGCGCGAGGGGACTATGTCATGTATGCTTAATGTGGAAGGCGGTGGTGGCTGAC 690
Db 559 GTGCGCGGTGGGACTATGTCAGGTTATGCTCAGGCTTATGCTCAGGCTGGAAGGCTGT 618
Qy 691 CGGGTTACCTGGAAGGCGCTGATATGTTCCGGGACGCTATTCATCTCCAGTCTTC 750
Db 619 AGCCCTACCTCCGCGAGGCGATGATGTTCCGGGACGCGACGAGGCCCGCTTTC 678
Qy 751 GTGTTTACAAGCAACGCTGCTGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAC 810
Db 679 GTGTCACCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
Qy 811 GTGCTGTTCCGGGCAATGTTATGAGGGTTCGCCAGCAAGGACTTCGGGCTCTCACC 870
Db 739 GTGACGTTTGTGCGGATGGACAGGAGCTACACCGTGAAGGCTTGGCTCTCTACA 798
Qy 871 CAGTGCACACACCATCATCATCTATTTGGGACCTTGGGATTTGGGCTGCTGCTGCTG 930
Db 799 CAGTGCACACACCATCATCATCTATTTGGGACCTTGGGATTTGGGCTGCTGCTGCT 858
Qy 931 GGTGTTGATACCATCTACTTAGCCAACTACACCTTCCGATTTCCGTTTCTCTCAAGTC 990
Db 859 GCGGAGACATGCTTCTACCTGCGCACTTACCTTCCAGACTCTGAGTCTCTGAAGATC 918
Qy 991 TTTAAGCCAGGAGGAGCTTCTTACCCGAAATGGGTGGGCACTCCCTGCGGATCTCC 1050
Db 919 TTTAAGCCAGGAGGAGCTTCTTCCCGGAGTGGGTGGGCAATTAATGACAGACTTCT 978
Qy 1051 CTC 1053
Db 979 CTC 981

RESULT 10
US-08-395-800A-9
; Sequence 9, Application US/08395800A
; Patent No. 5867732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J.
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; TITLE OF INVENTION: GENOTYPING A PERSON
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202


```

QY 1051 CTC 1053
Db 1589 CTC 1591

RESULT 14
US-08-273-411-2
; Sequence 2, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Fabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273.411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 104..1201
; OTHER INFORMATION: /note= "Nucleotides 104 through 1201 encode the GDP-L-fuc
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678.
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
US-08-273-411-2

Query Match 35.2%; Score 405; DB 1; Length 3373;
Best Local Similarity 68.4%; Pred. No. 3.7e-108;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 217 GGCATGTTCCAGCATCAATTCCTATGGCGCGGTGGGGAACCAAGATGGCGCAATACGCCACA 276
Db 338 GGCACCTGGACTGTCTACCCCAATGGCGGTGGTAATCAGATGGGACAGTATGCCAG 397

QY 277 CTCCTTGCACTGCCAGGATGAACGCGCTGGCTTCATCCCGCATCCATGCACAC 336
Db 398 CTCGTGGCTCTGCCAGCTCAACGCGCGCGGCGCTTATCTGCTGCCATGCGCC 457

QY 337 GCTCTAGCGCCCATCTTCAGGATCAGCCCTCCCGGTGTTACACAGCGACACGGCCAAAAG 396

```

```

Db 458 GCCCTGCCCGGTATTCCCGCATCACCTGCCCTGCCCTGGCCCGCCAGAGTGSACAGCGC 517
QY 397 ATCCCATGGCAGAAATTACCATCTCAAGACTGGATGGAGAGCGTTACCGGCACATCCG 456
Db 518 ACGCCGTGGCGGAGCTGCAGCTTCACGACTGGATGCGGAGGAGTACCGGAGCTTGAGA 577
QY 457 GGACACTTTGTGCGCTTCACGGGATACCCGTGCTCTGGACCTTCTACACACACTGCGC 516
Db 578 GATCCTTTCTTGAAGCTCTCTGGCTTCCCTGCTCTGGACTTCTTCCACCATCTCCGG 637
QY 517 CCAGAGATCTCTSAAGAGTTCCACCTGCATGACACACCTGCGGAGAGGCGCCAGCCCTC 576
Db 638 GAACAGATCCCGCAGAGTTCAOCCIGCACACACCTTGGGAAGAGGCGCAGAGTGTG 697
QY 577 CTGCGTGTCTGCGGGTG-----AATGGAGCAGCGGAGTACTTTTGTGGGTGTCCAT 630
Db 698 CTGGGTCAAGTCCCGCTGGCGGCACAGGGGACCGCCGCGACCTTTGTGCGCGTCCAC 757
QY 631 GTGCGCGCAGGGGACTATGTGCATCTCATGCTAATGTGTGAAGGGGTGCTGCTGAC 690
Db 758 GTGCGCGCTGGGACTATCTCAGGTTATGCTCAGCGCTGGAAGGGTGTGTTGGGGCAG 817
QY 691 CGGGTTACCTGGAAGGCGCTGCATATGTTCCGGGCACGCTATTCACTCCASTCTTC 750
Db 818 AGCGCTACCTCCGCGAGGCGCATGCTGTTCCGGGCACGCGACGAAAGCCCGCTTTC 877
QY 751 GTGTTACAAGCAACGGTATGGCTGTGCGGAGAGACATTAATGCTTCCCGAGGAGAC 810
Db 878 GTGGTCAACCAACGGCATGGAGTGTGTAAAGAAACATCGACACCTCCGAGGGCGAT 937
QY 811 GTGGTGTTCGCGGCAATGGTATTGAGGGGTGCGCAGCAAGGACTTCGCGCTGCTACC 870
Db 938 GTGACGTTTGTGGCATGGAGAGGCTACACGCTGGAAGAGACTTTGCCCTCTCACA 997
QY 871 CAGTGCAACCAACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCGCTACCTGGCA 930
Db 998 CAGTGCAACCAACCATCATGACTATTGGGACCTTTGGGCTGCGCTACCTGAGATC 1057
QY 931 GTGTGTATACCATCTACTTACCGCACTACACCTTCCGGATTCCTCCCTCAAGTGC 990
Db 1058 GCGGAGAGACTGTCTACCTGGCCAACTTACCTGCGGAGCTCTGAGTTCCTGAAGATC 1117
QY 991 TTAAAGCCAGAGGCGAGCTTCTACCGCAATGGTGGGATCCCTGCGGATCTGTCGCCA 1050
Db 1118 TTAAAGCCGAGGCGCGCTTCTGCGGAGTGGGTGGGCAATTAATGCAAGACTTGTCTCA 1177
QY 1051 CTC 1053
Db 1178 CTC 1180

RESULT 15
US-08-675-773B-3
; Sequence 3, Application US/08675773B
; Patent No. 6156288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; TITLE OF INVENTION: FOR XENOTRANSPLANTATION...
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 SEVENTH STREET, N.W., SUITE 300
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/675,773B
FILING DATE: 03-JUL-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: DIAMOND=1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3791 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-675-773B-3

Query Match 35.2%; Score 405; DB 4; Length 3791;
Best Local Similarity 68.4%; Pred. No. 3.9e-108;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY	217	GGCATGTTACAGATCAATTCATTCGCGGCTGGGAACAGATGGCGAATAGCCACACA	276
DB	2293	GGACCTGGACTGTCFACCCCAATGGCGGTTGGTAATCAGATGGGACAGATGCCACG	2352
QY	277	CTCTTTCGACTGGCCAGGATGACGGAGCGTTCGCTTCATCCCGCATCCATGCAACAAC	336
DB	2353	CTGCTGCTCTGGCCCACTCAACGGCGCGCGGCGCTTATCTGCTGCTGCATGCATGCC	2412
QY	337	GCCTAGCGCCCATCTTCAGATCAGCTCCGCTGTTTACACAGCGACACGGCCCAAAAAG	396
DB	2413	GCCCTGGCCCCGGTATTCGGCATCACCCTGCCCTGCTGGCCCCAGAGTGGACAGCGCG	2472
QY	397	ATCCCATGGCAGAAATACCAATCTCAAGACTGATGAGAGAGCGTTACCGCCACATCCG	456
DB	2473	ACGCGTGGCGGAGCTGCACCTTCAGACTGGATGTCGGAGGAGTACGGCGACTTGAGA	2532
QY	457	GGACACTTTGTGGCGTTCAAGGGATACCGTCTCTGGACCTTACACACCACTGGCG	516
DB	2533	GATCCCTTCCGGAAGCTCTCTGGCTCCCTGCTCTGGACTTCTTCCACCACTCCGG	2592
QY	517	CCAGAGTCCCTGAAGAGTTACCTCATGACCACTGCGGAGGAGGCGCCAGSCCTTC	576
DB	2593	GAACAGATCCCGAGAGATTCACCTGTCAGACCACTTCGGGAAGAGGCGCAGAGTGTG	2652
QY	577	CTGCTGCTCTGCGGGTG-----AATGGAGCCAGCCGAGTACTTTTGGGGTGTCCAT	630
DB	2653	CTGGGTACGCTCCGCTGGCGCGCACAGGGACCGCGCCGACCTTTGTCGGCGTCCAC	2712
QY	631	GTGCGCGGAGGGGATGATGTCATGCTCATGTTGTGGAAGGGGTGGTGGCTGAC	690
DB	2713	GTGCGCGGTGGGGACTATCTGAGGTTATGCTCAGCGCTGGAAGGGGTGGTGGCGAC	2772
QY	691	CGGGGTACTCTGGAAGAGCCCTGGATATGTTCCGGGCGACGCTATTCATCTCCAGTCTTC	750
DB	2773	AGCGCCCTACTCTCCGGCAGCGCATGAGTGTGTTCCGGGCGACGCGACAGACCCCGCTTC	2832
QY	751	GTGGTTACAGAACAGCTATGGCTGGTCCGGGAGAGAACTTAATGCTTCCCGAGGAGAC	810
DB	2833	GTGGTACACAGCAACGGCATGGAGTGTGTAAGAAACATCGACACCTCCCGAGGCGAT	2892
QY	811	GTGGTGTTCGGCGGCAATGATATGAGGGTCCCGACGACAGACTTCGGCTGCTCAAC	870
DB	2893	GTGACGTTTCTGGCGATGACAGAGGCTTACACCGTGGAAAGACTTGGCCCTGCTCACA	2952
QY	871	CAGTGAACCAACCATCATGACTATGGACCTTTGGGATTTGGGCTGCTACCTGGCA	930

Search completed: May 26, 2003, 11:58:32
Job time : 64.7104 secs

1	1149	100.0	1149	10	US-09-999-672-7	Sequence 7, Appl
2	1149	100.0	1149	12	US-10-040-863-7	Sequence 9, Appl
3	1068	93.0	1068	10	US-09-999-672-9	Sequence 9, Appl
4	1068	93.0	1068	12	US-10-040-863-9	Sequence 9, Appl
5	630.2	54.8	1043	10	US-09-051-034A-1	Sequence 1, Appl
6	632.6	53.3	3088	10	US-09-954-456-45	Sequence 45, Appl
7	632.6	53.3	3088	10	US-09-954-456-121	Sequence 1521, Ap
8	632.6	53.3	3088	10	US-09-969-347-234	Sequence 234, App
9	423.8	36.9	1098	10	US-09-051-034A-3	Sequence 3, Appl
10	423.8	36.9	1269	10	US-09-844-268-12	Sequence 12, Appl
11	423.8	36.9	1269	10	US-09-844-705-12	Sequence 12, Appl
12	405	35.2	3373	9	US-10-105-963-9	Sequence 9, Appl
13	405	35.2	8174	10	US-09-863-475A-5	Sequence 5, Appl
14	72.8	6.3	100	10	US-09-999-673-29	Sequence 29, Appl
15	72.8	6.3	100	12	US-10-040-863-29	Sequence 29, Appl
16	71.2	6.2	100	10	US-09-999-672-23	Sequence 23, Appl
17	71.2	6.2	100	10	US-09-999-672-26	Sequence 26, Appl
18	71.2	6.2	100	12	US-10-040-863-23	Sequence 23, Appl
19	71.2	6.2	100	12	US-10-040-863-26	Sequence 26, Appl

181 CGGACAGCAGCAGCATGGATGGAGAGCTGGGGCATGTCCAGATCAATTCATT 240
241 GGCGGCTGGGACCAAGATGGGAGATAGCCACACTCTTGCAGTGGCAGATGAAC 300
241 GGCGGCTGGGACCAAGATGGGAGATAGCCACACTCTTGCAGTGGCAGATGAAC 300
301 GGAGGCTTGGTTATCCCGGATCCATGACACAAAGCTCTAGCGCCCATCTTCAGGATC 360
301 GGAGGCTTGGTTATCCCGGATCCATGACACAAAGCTCTAGCGCCCATCTTCAGGATC 360
361 AGCCTCCCGGTGTACACAGCAGCAGCCGCAAAAGATCCCATGCGAGATTCATCTC 420
361 AGCCTCCCGGTGTACACAGCAGCAGCCGCAAAAGATCCCATGCGAGATTCATCTC 420
421 AACGACTGGATGGAGAGCGTTACGCGACATTCGGGACACTTGTGCGCTTCACGGGA 480
421 AACGACTGGATGGAGAGCGTTACGCGACATTCGGGACACTTGTGCGCTTCACGGGA 480
481 TACCGGTCTCTGGACCTTCTACACACACCTGCGCCGAGAGATCCTGAAGAGTTTACC 540
481 TACCGGTCTCTGGACCTTCTACACACACCTGCGCCGAGAGATCCTGAAGAGTTTACC 540
541 CTGATGACACAGCTGGGAGAGAGCCGAGCCCTTCTGCGTGTCTGCGGTGAATGGG 600
541 CTGATGACACAGCTGGGAGAGAGCCGAGCCCTTCTGCGTGTCTGCGGTGAATGGG 600
601 AGCCAGCGAGTACTTTTCTGGGTCTTACATGTGCGCGAGGACTATGTGATGTCTATG 660
601 AGCCAGCGAGTACTTTTCTGGGTCTTACATGTGCGCGAGGACTATGTGATGTCTATG 660
601 AGCCAGCGAGTACTTTTCTGGGTCTTACATGTGCGCGAGGACTATGTGATGTCTATG 660
661 CTTAATGTGTGAAGGCGGTGTGCGGTCTTACATGTGCGCGAGGACTATGTGATGTCTATG 720
661 CTTAATGTGTGAAGGCGGTGTGCGGTCTTACATGTGCGCGAGGACTATGTGATGTCTATG 720
721 TTCGGGCAAGCTATTCATCTCCAGTCTTCTGTTTACAAAGCAAGCGTATGCGCTGGTGC 780
721 TTCGGGCAAGCTATTCATCTCCAGTCTTCTGTTTACAAAGCAAGCGTATGCGCTGGTGC 780
781 CGGAGACATTAATGCTTCCGAGGAGACGTTGTTTCCGCGCAATGGTATTGAGGGG 840
781 CGGAGACATTAATGCTTCCGAGGAGACGTTGTTTCCGCGCAATGGTATTGAGGGG 840
841 TCGCCAGCAAGGACTTCCGCTGTCTACCCAGTGCACACACACCATCATGACTATTGGG 900
841 TCGCCAGCAAGGACTTCCGCTGTCTACCCAGTGCACACACACCATCATGACTATTGGG 900
901 ACCTTTGGGATTTGGGCTGCTTACCTGGCAGTGGTGTATACCATCTACTTACCCACTAC 960
901 ACCTTTGGGATTTGGGCTGCTTACCTGGCAGTGGTGTATACCATCTACTTACCCACTAC 960
961 ACCCTTCGGATTTCCGTTCTCAAAGTCTTAAAGCCAGAGGAGCCCTTCTACCCGAA 1020
961 ACCCTTCGGATTTCCGTTCTCAAAGTCTTAAAGCCAGAGGAGCCCTTCTACCCGAA 1020
1021 TGGGTGGGATCCCTGCGGATCTGTCGCCACTCTTAAAGGATTAACACAGGCTGTCT 1080
1021 TGGGTGGGATCCCTGCGGATCTGTCGCCACTCTTAAAGGATTAACACAGGCTGTCT 1080
1081 CGGTCCCATTCGACCTCAAGCAAAAGAGTACCTTGTACGTCCAGAGAGCGCTTC 1140
1081 CGGTCCCATTCGACCTCAAGCAAAAGAGTACCTTGTACGTCCAGAGAGCGCTTC 1140
1141 TGATGGGAA 1149
1141 TGATGGGAA 1149

RESULT 2
US-10-040-863-7
; Sequence 7, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:

APPLICANT: Eric H. Holmes et al.
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 8511-029
CURRENT APPLICATION NUMBER: US/10/040,863
CURRENT FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: 09/299,886
PRIOR FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1149
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1143)
US-10-040-863-7
Query Match 100.0%; Score 1149; DB 12; Length 1149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIGCCAGCGCCAGGTTCCCTTCTCTCTGCGCCACTTCTCTCATCTTCTCTTC 60
DB 1 ATGCGCAGCGCCAGGTTCCCTTCTCTCTGCGCCACTTCTCTCATCTTCTCTTC 60
QY 61 GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCA 120
DB 61 GTGACTTCCACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCA 120
QY 121 AAGGAATTTACCGATGACGACTCAAAATGTCTCTGGGAAACACAGAAAGCCAGAGT 180
DB 121 AAGGAATTTACCGATGACGACTCAAAATGTCTCTGGGAAACACAGAAAGCCAGAGT 180
QY 181 CGGACAGCGACGACGATGGGAATGAGAGCTGCGGGCATCTTCCAGATCAATTCATT 240
DB 181 CGGACAGCGACGACGATGGGAATGAGAGCTGCGGGCATCTTCCAGATCAATTCATT 240
QY 241 GCGCGCTGGGGAACAGATGGCGCAATACGCCACACTCTTTGCATGCGCCAGGATGAAC 300
DB 241 GCGCGCTGGGGAACAGATGGCGCAATACGCCACACTCTTTGCATGCGCCAGGATGAAC 300
QY 301 GAGCGCTTGGTTCATCCCGCATCCATGCACAGGCTCTAGCGCCCATCTTCAGGATC 360
DB 301 GAGCGCTTGGTTCATCCCGCATCCATGCACAGGCTCTAGCGCCCATCTTCAGGATC 360
QY 361 AGCCTCCCGGTGTACACAGCGACAGCGCCAAAGATCCCATGGCAGAAATACCATCTC 420
DB 361 AGCCTCCCGGTGTACACAGCGACAGCGCCAAAGATCCCATGGCAGAAATACCATCTC 420
QY 421 AACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGACACTTTGTGCGGTTCAGGGA 480
DB 421 AACGACTGGATGGAGAGCGTTACCGCCACATTCGCGGACACTTTGTGCGGTTCAGGGA 480
QY 481 TACCGGTCTCTGGAGCTTACACACACTTCCGCGCCAGAGATCTCTGAGGAGTTTCAAC 540
DB 481 TACCGGTCTCTGGAGCTTACACACACTTCCGCGCCAGAGATCTCTGAGGAGTTTCAAC 540
QY 541 CTGCAATGACCGTGGGAGAGGCGCCAGGCTTCTCGTGTGTCTGCGGTGAATGGG 600
DB 541 CTGCAATGACCGTGGGAGAGGCGCCAGGCTTCTCGTGTGTCTGCGGTGAATGGG 600
QY 601 AGCCAGCGAGTACTTTTGTGGTGTCCATGTGCGCGGAGGAGCTATGTGATGTCTATG 660
DB 601 AGCCAGCGAGTACTTTTGTGGTGTCCATGTGCGCGGAGGAGCTATGTGATGTCTATG 660
QY 661 CTTAATGTGTGAAGGCGGTGTGCGGTCTTACCTGAGAGGCGCTTGGATATG 720
DB 661 CTTAATGTGTGAAGGCGGTGTGCGGTCTTACCTGAGAGGCGCTTGGATATG 720
QY 721 TTCGGGCAAGCTATTCATCTCCAGTCTTCTGTTTACAAAGCAAGCGTATGCGCTGGTGC 780


```
Db 721 TCCGGGACACGTAATTCATCTCCAGCTTCGTGGTTACAGCAAGGTAATGCCCTGGTGC 780
QY 781 CGGAGAACATTAATCCCTCCCGAGGAGACGTGGTTCGGGGCAATGGTAATGAGGG 840
Db 781 CGGAGAACATTAATCCCTCCCGAGGAGACGTGGTTCGGGGCAATGGTAATGAGGG 840
QY 841 TCGCCAGCAAGGACTTCGGCGTGTCTACCCAGTGCACACCAACCATCATGACTATTGG 900
Db 841 TCGCCAGCAAGGACTTCGGCGTGTCTACCCAGTGCACACCAACCATCATGACTATTGG 900
QY 901 ACCCTTCGGGATTCGGCTGCCTACCTGGCAGTGGTGATACCATCTACTTACCAACTAC 960
Db 901 ACCCTTCGGGATTCGGCTGCCTACCTGGCAGTGGTGATACCATCTACTTACCAACTAC 960
QY 961 ACCCTTCGGGATTCGGCTGCCTACCTGGCAGTGGTGATACCATCTACTTACCAACTAC 1020
Db 961 ACCCTTCGGGATTCGGCTGCCTACCTGGCAGTGGTGATACCATCTACTTACCAACTAC 1020
QY 1021 TGGGTGGGATTCCTCCGATCTGTCCTCCAGTGCCTTAAAGCCAGAGCGCTTCTACCCGAA 1080
Db 1021 TGGGTGGGATTCCTCCGATCTGTCCTCCAGTGCCTTAAAGCCAGAGCGCTTCTACCCGAA 1080
QY 1081 CGGTCCCACTTCCACTCAAGCAAGGAGTCACTTGTACGTGCAGGAAGAGGCTTC 1140
Db 1081 CGGTCCCACTTCCACTCAAGCAAGGAGTCACTTGTACGTGCAGGAAGAGGCTTC 1140
QY 1141 TGATGGGAA 1149
Db 1141 TGATGGGAA 1149

RESULT 3
US-09-999-672-9
; Sequence 9, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
US-09-999-672-9

Query Match 93.0%; Score 1068; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CTCACGACGCAATAGTAGCTCCACCCCTGTCCAGAGGAATTAACCGATCAGCACT 141
Db 1 CTCACGACGCAATAGTAGCTCCACCCCTGTCCAGAGGAATTAACCGATCAGCACT 60
QY 142 CAATGTCTCGGGAACACAGAAAGCCAGAGATCGACGGGACAGCAGCATGGG 201
Db 61 CAATGTCTCGGGAACACAGAAAGCCAGAGATCGACGGGACAGCAGCATGGG 120
QY 202 AATGAGAGCTGCGGGGATGTTTCCAGATCAATTCATTTGGCGGGCTGGGGAACAGATG 261
Db 121 AATGAGAGCTGCGGGGATGTTTCCAGATCAATTCATTTGGCGGGCTGGGGAACAGATG 180
```

```
QY 262 GCGCAATAGCCACACTCTTTGCACTGGCCAGGATCAAGGAGGCTTCGCTTCATCCCC 321
Db 181 GCGCAATAGCCACACTCTTTGCACTGGCCAGGATCAAGGAGGCTTCGCTTCATCCCC 240
QY 322 GCATCCATGCACAAAGCTCTAGCGCCATCTTTCAGGATCAGCTCCCGGTGTACACAG 381
Db 241 GCATCCATGCACAAAGCTCTAGCGCCATCTTTCAGGATCAGCTCCCGGTGTACACAG 300
QY 382 GACAGGGCCAAAGATCCCATGGGAGAAATACCATCTCAAGACTGGATGGAGGAGCGT 441
Db 301 GACAGGGCCAAAGATCCCATGGGAGAAATACCATCTCAAGACTGGATGGAGGAGCGT 360
QY 442 TAGCCGCCAATTCGGGACACTTTGCGCTTCACGGGATACCGGTCTCTCGGACCTTC 501
Db 361 TAGCCGCCAATTCGGGACACTTTGCGCTTCACGGGATACCGGTCTCTCGGACCTTC 420
QY 502 TACCACCACTGCGCCAGAGATCTTGAAGGAGTTCACCTGTCATGACACGTCGCGGAG 561
Db 421 TACCACCACTGCGCCAGAGATCTTGAAGGAGTTCACCTGTCATGACACGTCGCGGAG 480
QY 562 GAGGCCAGGCTTCCTGCGTGGTCTGGGGTGAATGGGAGCCAGCGGAGTCTTTGTG 621
Db 481 GAGGCCAGGCTTCCTGCGTGGTCTGGGGTGAATGGGAGCCAGCGGAGTCTTTGTG 540
QY 622 GGTGTCCATGTGGCGGAGGGGACTATGTGCTATGCTATGCTGTAAGGGCGTG 681
Db 541 GGTGTCCATGTGGCGGAGGGGACTATGTGCTATGCTATGCTGTAAGGGCGTG 600
QY 682 GTGCTGACCGGGGTTACCTGGAAAGGCCCTGGATATGTCGGGACGCTATTCATCT 741
Db 601 GTGCTGACCGGGGTTACCTGGAAAGGCCCTGGATATGTCGGGACGCTATTCATCT 660
QY 742 CCAGTCTTCGTGTACAGCAACGCTATGCTGCTGGCGGAGACATTAATGCTTC 801
Db 661 CCAGTCTTCGTGTACAGCAACGCTATGCTGCTGGCGGAGACATTAATGCTTC 720
QY 802 CGAGGAGACGTGTGTTCGGGGCAATGTTATGAGGGTTCGCGAGCAAGGACTTCGCG 861
Db 721 CGAGGAGACGTGTGTTCGGGGCAATGTTATGAGGGTTCGCGAGCAAGGACTTCGCG 780
QY 862 CTGCTCACCAGTGCACACCAACCATCATGACATATTTGGGACCTTTGGGATTTGGCTGCC 921
Db 781 CTGCTCACCAGTGCACACCAACCATCATGACATATTTGGGACCTTTGGGATTTGGCTGCC 840
QY 922 TACCTGCGAGGTGGTATACCATCTACTTACGCAACTACACCTTCGCGGATTCCTCGTTC 981
Db 841 TACCTGCGAGGTGGTATACCATCTACTTACGCAACTACACCTTCGCGGATTCCTCGTTC 900
QY 982 CTCAAAGTCTTTAAGCCAGAGGAGCGCTTCCTACCCGAATGGGTGGCATCCCTGCGGAT 1041
Db 901 CTCAAAGTCTTTAAGCCAGAGGAGCGCTTCCTACCCGAATGGGTGGCATCCCTGCGGAT 960
QY 1042 CTGTCGCCACTCTTAAAGGATTAACACGACGCTGCTCTCGGTCCCACTTCCACTCAAG 1101
Db 961 CTGTCGCCACTCTTAAAGGATTAACACGACGCTGCTCTCGGTCCCACTTCCACTCAAG 1020
QY 1102 GCAAAAGGAGTCACTTGTAGTGCAGAGGAGCGCTTCTGATGGGA 1149
Db 1021 GCAAAAGGAGTCACTTGTAGTGCAGAGGAGCGCTTCTGATGGGA 1068

RESULT 4
US-10-040-863-9
; Sequence 9, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
```

```
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
US-10-040-863-9

Query Match      93.0%; Score 1068; DB 12; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  82 CTCAGCAGCGAATAGTAGCTCCAAACCCCTGTGAGAGAGGAATTAACGATGACGACT 141
Db  1 CTCAGCAGCGAATAGTAGCTCCAAACCCCTGTGAGAGAGGAATTAACGATGACGACT 60

Qy  142 CAATGTCCTCGGGAACACAGAAAGCCAGAGATGCGCGGACAGGACGACGATGGG 201
Db  61 CAATGTCCTCGGGAACACAGAAAGCCAGAGATGCGCGGACAGGACGACGATGGG 120

Qy  202 AATGAGAGCTGCGGGCATGTTTCCAGATCAATTCATTTGCGCGCTGGGGAACCAATG 261
Db  121 AATGAGAGCTGCGGGCATGTTTCCAGATCAATTCATTTGCGCGCTGGGGAACCAATG 180

Qy  262 GCGGAATAGCCACACTCTTTGCACTGCGCAGGATGAACGACGCGTGTGCTTCATCCC 321
Db  181 GCGGAATAGCCACACTCTTTGCACTGCGCAGGATGAACGACGCGTGTGCTTCATCCC 240

Qy  322 GCATCCATGCAACAGCTTAGCGCCCATCTTCAGGATCAGCCTCCGCGTGTACACAGC 381
Db  241 GCATCCATGCAACAGCTTAGCGCCCATCTTCAGGATCAGCCTCCGCGTGTACACAGC 300

Qy  382 GACACGCGCCAAAAGATCCCATGTCAGAAATTAACATCTCAACGACTGGATGGAGAGCGT 441
Db  301 GACACGCGCCAAAAGATCCCATGTCAGAAATTAACATCTCAACGACTGGATGGAGAGCGT 360

Qy  442 TACCGCACATTCGCGGACACTTTGTGGGCTTACGGGATACCGGTGCTCCTGACCTTC 501
Db  361 TACCGCACATTCGCGGACACTTTGTGGGCTTACGGGATACCGGTGCTCCTGACCTTC 420

Qy  502 TACCACACCTGCGCCAGAGATCTGGAAGAGTTCACCTGATGACACGTCGCGGAG 561
Db  421 TACCACACCTGCGCCAGAGATCTGGAAGAGTTCACCTGATGACACGTCGCGGAG 480

Qy  562 GAGGCCAGGCGCTTCTCGCGTGTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTG 621
Db  481 GAGGCCAGGCGCTTCTCGCGTGTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGTG 540

Qy  622 GGTGTCCATGTGCGCCGAGGAGTATGTGCAITCATGCCATATGTGTGGAAGGCGTG 681
Db  541 GGTGTCCATGTGCGCCGAGGAGTATGTGCAITCATGCCATATGTGTGGAAGGCGTG 600

Qy  682 GTGGTGACCGGGTTACTTGGAAAAGCCCTGGATATGTTCCGGGACCGCTATTTCATCT 741
Db  601 GTGGTGACCGGGTTACTTGGAAAAGCCCTGGATATGTTCCGGGACCGCTATTTCATCT 660

Qy  742 CCAGTCTTCTGGTGTACAAAGCAACGGTATGCGCTGGTGGCGGGAGAACATTAATGCTTCC 801
Db  661 CCAGTCTTCTGGTGTACAAAGCAACGGTATGCGCTGGTGGCGGGAGAACATTAATGCTTCC 720

Qy  802 CGAGGAGAGCTGCTGTTCCCGGGCAATGGTATTGAGGGGTGCGCCAGGACTTCGCG 861
Db  721 CGAGGAGAGCTGCTGTTCCCGGGCAATGGTATTGAGGGGTGCGCCAGGACTTCGCG 780

Qy  862 CTGCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 921
Db  781 CTGCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 840

; PRIOR APPLICATION NUMBER: 09/051,034A-1
; Sequence 1, Application US/09051034A
; Patent No. US20010055584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051,034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1043
; TYPE: DNA
; ORGANISM: Sus Domesticus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1031)
US-09-051-034A-1

Query Match      54.8%; Score 630.2; DB 10; Length 1043;
Best Local Similarity 76.9%; Pred. No. 2.2e-191;
Matches 814; Conservative 0; Mismatches 203; Indels 42; Gaps 2;

Qy  1 ATGCCAGCGCGCCAGGTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db  9 ATGTCAGCATGCGAGGCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 68

Qy  61 GTGACTTCCACCATATCCACCTCCAGCAGGAGATGAGAGCTCCACACCCCTGTGACAG 120
Db  69 ACGGCTTCCACCATATTTCACTTTCAGCAGAGGATGAGAGATTCACCCACCTG----- 123

Qy  121 AAGGAATTAACGATGACGACTCAATGTCCTGGGAACACAGAAAGCCAGAGATGCGA 180
Db  124 -GGGAGTTACAGATGGTGACCGAGTGACACACAGAGCCCTCGAGCCC----- 172

Qy  181 CGGACACGCGAGCAGCATGGGAATGGAGAGCTGGGGGSCATGTTACAGATCAATTCAT 240
Db  173 -----CCAGCTGAAGGSCATGTGGACCATCAATGCCATC 206

Qy  241 GCGCGGCTGGGGAACACAGATGGGGAATACCCACACTCTTTGGCACTGGCCAGGATGAAC 300
Db  207 GCGCGCTGGGGAACACAGATGGGGAATACCCACCTGTACGGGCTGCCAGGATGAAC 266

Qy  301 GGAGGCTGTGGTTCATCCCGGCATCCATGACACAGCTCTAGGCGCCCACTCTCAGGATC 360
Db  267 GGGGGGCGGCGCTTCATCCCGCCCGAGATGCACAGCGCTGGGCCCCCATCTTCAGGATC 326
```


Db 867 CATTGACACCTCCACCGTGTATGTTGGTGGGATGGCATTTGAGGCTCACCTGC 926
Qy 849 CARGACTTCGGTGTCTACCCAGTTCACACACACACACATCATGACTATTGGACCTTTGG 908
Db 927 CAAGAATTTTGTCTACTACACAGTGTAAACACACACATCATGACCATTTGGAGCTTCGG 986
Qy 909 GATTGGGCTGCCTTACCTGCGAGTGTGTATACCATCTACTATTAGCCAACTACACCTTCC 968
Db 987 GATCTGGGCGGATCTACGCGGAGACACCATCTACTGSCCAATTACACCTTCC 1046
Qy 969 GGATTTCCGTTCTCTCAAGTCTTTAAGCCAGAGGAGCCTTCTTACCCGAATGGGTGG 1028
Db 1047 CGACTCCCTTTCTCTCAAAATCTTTAAGCCAGAGGAGCCTTCTGCGGAGTGGACAG 1106
Qy 1029 CATCCCTGCGCATGTGTCACACCTTAAAGCATTTAACCACGCTGTCTCTGGTCCCA 1088
Db 1107 GATTGCGGAGAGCTGTGCCCCCTTACTCAAGCACTAATGCTGGCCCATCTTTGAGACCT 1166
Qy 1089 CTTCCACCTCAAGGC 1103
Db 1167 TTTCTCTCTCTGTC 1181

RESULT 7

US-09-954-456-1621
; Sequence 1621, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Candi
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR FILING DATE: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1621
; LENGTH: 3088
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1621

Query Match 53.3%; Score 612.6; DB 10; Length 3088;
Best Local Similarity 73.8%; Pred. No. 1.5e-185;
Matches 808; Conservative 0; Mismatches 254; Indels 33; Gaps 1;
Qy 9 CGCCAGGTTCTTCTTCTCTTCTCTGCGCCACTTCTCTATCTTCTCTGCTGACTTC 68
Db 120 CGTTCAGATGCTTCTCTCTTCTCTTCTCTGCGCCACTTCTCTTCTCTTCTCTGCTGACTTC 179
Qy 69 CACCATCATCCACTCCAGCAGGCAATGAGTGAAGCTCCACCCCTGTACAGAGGAAT 128

Db 180 CACTATATTTTACGCTTCAGCAGCGGCTAGCGAAGATTCAAGCCATGTGGGAGTTACCGGT 235
Qy 129 ACCGATGACGACTCAATGCTCTCGGGAACACACAGAAGCCAGAGATGCCACGGGACAG 188
Db 240 GCAGATACCACTAGTGTAGCTCAACATCAAGAGCACTGGGACCCA----- 283
Qy 189 CGAGCAGCATGGGAATGGAGAGTTCGCGGGGATTTTCCAGCATCAATTCATTTGGCCCGCT 248
Db 284 -----GCCAGCTCAGGGGATGTGGACGATCAATGCAATAGCCCGCT 326
Qy 249 GGGGAACACAGTGGGCAATACGCCACACACTTTTGGCACTGGCCAGGATGAACGACGCT 308
Db 327 GGGGAACACAGTGGGCAATACGCCACACACTTTTGGCACTGGCCAGGATGAACGACGCT 386
Qy 309 TGCCTTCATCCCGCATCTTACGCAACAGCTCTAGCGCCATCTTTCAGGATCAACCTTCC 368
Db 387 CGCTTCATCCCGCATCTTACGCAACAGCTCTAGCGCCATCTTTCAGGATCAACCTTCC 446
Qy 369 GGTGTTACACAGCAGCAGCGCAAAAGATCCCATGGCAGAAATTACCATCTCAACGACTG 428
Db 447 GGTGTTACACAGCAGCAGCGCAAAAGATCCCATGGCAGAAATTACCATCTCAACGACTG 506
Qy 429 GATGAGGAGCGTTACCGCCACACTTCCGGGACACTTTTGGCGCTTCACGGGATACCGGTG 488
Db 507 GATGAGGAGCGTTACCGCCACACTTCCGGGAGGATGATGCTGCTTCCCGGCTTACCGCTG 566
Qy 489 CTCCTGGACCTTCTACACACACTTCCCGCCAGAGATCTTGAAGGATTCACCTTGCATGA 548
Db 567 CTCTGGACCTTCTACACACACTTCCCGCCAGAGATCTTCCAGGATTCACCTTGCATGA 526
Qy 549 CCAGTTCGGGAGGAGCGCCAGCGCTTCTGCGGTGCTGCGGTGTAATGGAGCCAGCC 608
Db 627 CCAGTTCGGGAGGAGCGCCAGCGCTTCTGCGGTGCTGCGGTGTAATGGAGCCAGCC 686
Qy 609 GAGTACTTTTCTGGTCTCCATGTGCGCGGAGGACTGTCATCTCATCTGCTTATGT 668
Db 687 GGGCAGCTTTGTAGGGTCTCATGTTCGCGGAGGACTGTCATCTCATCTGCTTATGT 746
Qy 669 GTGAAGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 728
Db 747 GTGAAGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 806
Qy 729 AGGATATCTCTCAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 788
Db 807 TCGTACAGCTTCT 866
Qy 789 CATTATGCTTCCGAGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
Db 867 CATTGACACCTTCCCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 926
Qy 849 CAAGGACTTCGGCTGCTCACCAGTGCACACACACATCATGACTATTGGACCTTTGG 908
Db 927 CAAAGATTTTGTCTTACTACACAGTGTAAACACACATCATGACCATTTGGAGCTTCGG 986
Qy 909 GATTGGGCTGCTTACCTGCGAGTGTGTATACCATCTTCTAGCCAACTACACCTTCC 968
Db 987 GATCTGGGCGCATACCTCAGCGGAGAGACACATCTACTCTGCGCAATTACACCTTCC 1046
Qy 969 GGATTTCTCCGTTCTCTCAAGTCTTTAAGCCAGAGGAGCCTTCTTACCCGAATGGGTGG 1028
Db 1047 CGACTCCCTTTCTCTCAAAATCTTTAAGCCAGAGGAGCCTTCTTCCCGGAGTGGACAG 1106
Qy 1029 CATCCCTGCGCATGTGTCGCCACTCTTAAAGGCAATTAACACAGCCTGTCTCTGGTCCCA 1088
Db 1107 GATTGCGCAGAGCTGTGCGCCCTTACTCAAGCACTAATGCTGGCCCATTTCTTTCAGACCT 1166
Qy 1089 CTTCCACCTCAAGGC 1103
Db 1167 TTTCTCTCTCTGTC 1181

RESULT 8
US-09-969-347-234

; Sequence 234, Application US/09969347
; Patent No. US2002011508A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 234
; LENGTH: 3088
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-234

Query Match 53.3%; Score 612.6; DB 10; Length 3088;
Best Local Similarity 73.8%; Pred. No. 1.5e-185;
Matches 808; Conservative 0; Mismatches 254; Indels 33; Gaps 1;

QY 9 CGCCAGGTTCCCTTCTCTCCCTTCCCTGTGGCCACCTTCCTCATCTTTGTCTTGGTGCATTC 68
Db 120 CGTTGAGATGCCCTTCTCTCTTCCCATGCCCCCATTCATCTCTTTGTCTTACGGTTTC 179

QY 69 CACCATATCCACCTCCAGCAGCGAATAGTGAAGCTCCCAACCCCTGTGAGAGGAAT 128
Db 180 CACTATATTTCAGTTTCAGCAGCGGTACGAGATTCAGCCATGCGGAGTTACCGGT 239

QY 129 ACCGATGAGCATCAAAATGTCTCGGGAACACAGAAAGCCAGAGATCGAGCGGACAG 188
Db 240 GCAGATACCAAGTGTAGCTCAACATCAAGGCACTGGGACCA----- 283

QY 189 CGAGCAGCATGGGAATGGAGAGCTGGGGCATGTTCAGATCAATTCATTCGGCCGGCT 248
Db 284 -----GCCAGCTAGGGGGATGTGGACGATCAATGCAATAGGCGCCT 326

QY 249 GGGGAACAGATGGGGAATAGCCACACTTTGCACTGGCCAGATGAACGAGCGGT 308
Db 327 GGGGAACAGATGGGGAATAGCCACACTTACGCCCTGGCCAGATGAACGAGCGGC 386

QY 309 TCGTTATCCCGCATCCATGCACAAACGCTTAGCGCCCATCTTCAGGATCAGCCTCCC 368
Db 387 CGCCTTCATCCCGGCCAGATGCACAGCACCTCGGCCCATCTTCAGAATCACCTGCC 446

QY 369 GTGTGTACAGCAGCACGCCCAAAAGATCCATGCGAGATTAACATCTCAACGACTG 428
Db 447 GTGTGTACAGCAGCACGCCCAAGATCCCTGGCAGATCCCTGGCAGACTACCATCTGAACGACTG 506

QY 429 GATGAGGAGCGTTTACCGCCCATTCGCGGACACTTTGTGCGCTTCACGGGATACCGCTG 488
Db 507 GATGAGGAGGAAATACCGCCACTTCCTCGGGGAGTACGTCCTTCACCGGCTACCCCTG 566

QY 489 CTCCTGGACCTTCTACCAACACCTGGGCCAGAGATCCCTGAAGAGGTTCAACCTGCATGA 548
Db 567 CTCCTGGACCTTCTACCAACACCTCGCCAGAGATCCCTGAAGAGGTTCAACCTGCATGA 626

QY 549 CCACGTGGGGAGGAGCGCCCTTCCTCGGTGCTCGCGGTGAATGGGAGCGACGC 608
Db 627 CCACGTGGGGAGGAGCGCCCAAGTTCTCGGGCCCTGCGAGGTGAACGGGAGCGCGCC 686

QY 609 GAGTACTTTTGTGGTGTCCATGTGCGCGAGGGACTATGTGCATGTCTCATGCTCAATGT 668
Db 687 GGGCACTTTGTAGGGTCCATGTTGCGCGAGGGACTATGTCCATGTCTATGCCAAAGT 746

QY 669 GTGGAAAGCGCTGTGGCTGACCGGGTTACTGTGAAAGAGGCCCTTGATATGTTCCGGGC 728
Db 747 GTGGAAAGGGGTGGTGGCGGACCGGCGATACCTACAGCAGGCGCTGGACTGTTCCGAGC 806

QY 729 ACSCATTAATTCATCCAGTCTTCTGTTTACAGAACGAGGTATGGCCTGTGTCGGGAGAA 788
Db 807 TCCTACAGCTCCCTCATCTTCTGTTGTCACCAAGTAAATGGATGGCCTGTGTCGGGAGAA 856

QY 789 CATTAATGTTCCCGAGGAGACCTGTGTTTCGGGGCAATGGTATTGAGGGGTGCGCCAGC 848
Db 867 CATGACACTCCACAGGGTGTGTTTGTGGGATGGCATGAGGGCTCACCTGC 926

QY 849 CAAGGACTTGGCGCTGCTCACCCAGTCAACACACCATCATGACTATTGGGACCTTTGG 908
Db 927 CAAAGATTTTGTCTACTACACAGTGAACACACCATCATGACTTGGGACGTTTCGG 986

QY 909 GATTTGGGCTGCTTACTGCGCAGGTGGTGTATACATCTACTTAGCCAACTACACCCCTTC 968
Db 987 GATCTGGCGGCATACCTCACGGCGGAGACACATCTACCTGGCCAAATACACCCCTCC 1046

QY 969 GGATTTCTCGTTCCTCAAAAGTCTTAAAGCCAGAGGAGCCTTCTACCCCAATGGGTGG 1028
Db 1047 CGACTCCCTTCTCTCAAAATCTTAAAGCCAGAGGAGCCTTCTGCGGAGTGGACAGG 1106

QY 1029 CATCCTGCGGATCTGTCCCACTCTTAAAGCATTAACACAGCCTGTCTCGGTCCCA 1088
Db 1107 GATTCGGCAGAGCTGTCCCTTACTCAAGCATAATGCTGCGCCATCTCTTGAGACCT 1166

QY 1089 CTTCCACCTCARGGC 1103
Db 1167 TTTCTCTTCTCTGTC 1181

RESULT 9
US-09-051-034A-3
; Sequence 3, Application US/09051034A
; Patent No. US2001005584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; TITLE OF INVENTION: GLYCOSYLTRANSFERASE
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051,034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Sus Domesticus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1098)
US-09-051-034A-3

Query Match 36.9%; Score 423.8; DB 10; Length 1098;
Best Local Similarity 69.6%; Pred. No. 2.6e-125;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

QY 217 GGCATGTTCCAGATCAATTCCTATTTGGCCGGTGGGGAACACAGATGGCGGAATACGCCACA 276
Db 235 GGGACCTTGACTATTTACCCGGATGGCCGGTTTGGGAACAGATGGGAGATGACGAGG 294

QY 277 CTCCTTGGCACTGGCCAGGATGAACGAGCTTGCCTTCATCCCCCATCCATGCAACAC 336
Db 295 CTCGTGGCCCTGGCGCAGCTCAACGGCCGCCAGGCGCTTCATCCAGCTGCCATGACGCC 354

QY 337 GCTGTAGCGCCCATCTTTCAGGATCAACCTCCCGGTGTTACACAGGACACGCGCAAAAG 396
Db 355 GTCTTGGCCCCCGTGTTCGCGATCAGCTGCTGCTGCTGGCGCCGAGGTAGACAGGCAC 414


```

: APPLICANT: VOGELI, PETER
:
: TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
:
: TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES.
:
: FILE REFERENCE: 21419/90369
:
: CURRENT APPLICATION NUMBER: US/09/844,705
:
: PRIOR FILING DATE: 2001-04-27
:
: PRIOR APPLICATION NUMBER: 09/443,766
:
: PRIOR FILING DATE: 1999-11-19
:
: NUMBER OF SEQ ID NOS: 13
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 12
:
: LENGTH: 1269
:
: TYPE: DNA
:
: ORGANISM: Porcine
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (9)..(1103)
:
: US-09-844-705-12

```

	Query Match	36.9%;	Score 423.8;	DB 10;	Length 1269;
	Best Local Similarity	69.6%;	Pred. No. 2.8e-125;		
	Matches 591;	Conservative 0;	Mismatches 252;	Indels 6;	Gaps 17;
QY	217	GGCATGTTACGATCAATTCATTTGGCCGGCTGGGGAACACAGATGGGCGAATACGCCACA	276		
Db	243	GGACCTGGACTATTTACCGCGATGGCCGGTTTGGGAACACAGATGGGACAGTATGCCACG	302		
QY	277	CTCTTTGCACCTGCCAGGATGAAGGACGGCTTCGTTTCATTCCTCCCGCATCCATGCAACAAC	336		
Db	303	CTGCTGGCCCTGGCGAGCTAAAGGGCGGCCAGGCCCTTCATCCAGCTGGCATGCAACGCCC	362		
QY	337	GCCTTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTATACACGACACAGGCCCAAAAG	396		
Db	363	GTCCGTGGCCCCGTGTCGCATACAGCTGCCTGTCTCTGGCGCCGAGGTAGACAGGCAC	422		
QY	397	ATCCCATGCGCAATTTACCATCTCAAGCATCTGATGAGGAGGAGGTATACCGCCACATTCG	456		
Db	423	GCCTCTTGGCGGAGCTTGGAGCTTCAGACTGGATGTCCGAGGATTTATGCCACTTAAAG	482		
QY	457	GGACACTTTTGTGGCTTCADGGGATACCGTGTCTGTCCCTGGACCTTCTPACCAACCACTGGC	516		
Db	483	GAGCCCTGGCTGAAGCTCACGGGCTTCCCTGTCTGTGGACCTTCTPCCACCACTCCGG	542		
QY	517	CCAGAGATCCTGAAGAGTTACCCCTGCATGACACAGTCCGGGAGGAGGCCAGGCCCTTC	576		
Db	543	GAGCAGATCCGACGAGTTACCCCTGCAGACCACTTCGCGCAGAGGCCAGGGGGTA	602		
QY	577	CTGCGT-----GGTCTGGSGTGATGGAGGCCAGCGAGTACTTTTGTGGGTGTCCAT	630		
Db	603	CTGAGTCACTTCGCTCTACCCCGACAGGGGACGCCGCCACGACCTCTGTGGGGGTCCAC	662		
QY	631	GTGCGCCGAGGGGACTATGTGCATGTCAATGTCTATGTTGTGGAAGGCGGTGGTGGCTGAC	690		
Db	663	GTGCGCCGCGGGACTATCTGGTGTATGATGCCCAAGCGCTGGAGAGGGGTGTGGGTGAC	722		
QY	691	CGGGGTTACTGTGAAAGGCCCTGATATGTTCCGGCCAGCTATTATCTCCAGCTCTTC	750		
Db	723	GGCCGTACTCTCAGCAGGCTATGGACTGGTTCGGGGCCGATACGAAGCCCCGCTTT	782		
QY	751	GTGGTTACAGCAACGGTATGSCCTGGTGC CGGGAGAACATTAATGCTTCCCGAGGAGAC	810		
Db	783	GTGGTCAACGACACGGCATGGAGTGGTCCCGGAAGAACATCGACACCTCCCGGGGGGAC	842		
QY	811	GTGGGTTCCGGGGCAATGGTATTGAGGGGTGCCACGCAAGGACTTCGGGCTGCTCAAC	870		
Db	843	GTGATCTTTCITGGCGATGGCGGGAGGCCGCCGCCAGGAGACTTGGCTGTCTGGTG	902		
QY	871	CAGTGCACCAACCAATCATGACTATTGGGAACCTTTGGGATTTGGGCTGCTACCTGGCA	930		
Db	903	CAGTGCACCAACCAATCATGACTATTGGGATTTGGGCTTCTGGGCGGCTACCTGGCT	962		
QY	931	GTGTGGTATACCATCTACTTAGCCCACTACACCTTCCCGGATTTCTCCGTTCTCTCAAGTC	990		

```

Db      963  GGTGAGATACCATCTACTTGGCTAACTTCACCTGGCCACTTCACGCTTCCTGAAGATC 1022
Qy      991  TTTAAGCCAGAGGAGCGCTTCTACTACCGCAATGGGTGGGCAATCCCTGCGCATCTGTCCCCA 1050
Db      1023  TTTAAACCCGAGGCTGCTTCTCTCCCGAGTGGGTGGCAATTAATGACAGACTGTCTCCA 1082
Qy      1051  CTCCTTTAAG 1059
Db      1083  CTCGAGATG 1091

RESULT 12
US-10-963-9
; Sequence 9, Application US/10105963
; Publication No. US20030068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection Syst
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 3373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)..(1201)
; OTHER INFORMATION:
; US-10-105-963-9

```

Query Match	35.2%;	Score 405;	DB 9;	Length 3373;
Best Local Similarity	58.4%;	Pred. No. 4.3e-119;		
Matches 577;	Conservative 0;	Mismatches 260;	Indels 6;	Gaps 1;
QY	217	GGCATGTTCACGATCAATTCATATGGCGCGCTGGGAACACAGATGGCGGATACGCCACA	276	
Db	338	GGCACCTGGACTGCTACCCCAATGGCCGGTTGGTAAATCAGATGGACAGTATGCCACG	397	
QY	277	CTCTTTGCATGGCAGGAGTAACGGACGGCTGCTCGTTCAATCCCGGATCCATGACACAAC	336	
Db	398	CTGCTGGCTCTGCCGCCAGCTCAACGGCGCGCGGCCCTTATCTCTGCTCGCATCATGCC	457	
QY	337	GCCTTAGCGCCCATCTTACGGATCAGCCTCCCGGTGTATACAGCGACACGGCCAAAAG	396	
Db	458	GCCTGGCCCCGTATATCCGCATCACCTCCCGTCTGGCCCCAGAGTGGACAGCCGC	517	
QY	397	ATCCATGGCAGAAATPACCAATCTACAGACTGATGGAGGAGCGTTTACCGCCACATTCOG	456	
Db	518	ACCCGCTGGCGGAGCTGTCAGCTTCAAGACTGGATGTCGGAGGAGTACGGGACTTGAGA	577	
QY	457	GGACATTTCTGCGCTTCACGGGATACCCGTGCTCCTGGACCTTCTACGACCACTGCGC	516	
Db	578	GATCCTTTCTGAAGCTCTGCGTTCCTCCCTGCTCTGGAGCTTTCTTCCACCACTCCGG	637	
QY	517	CCAGAGATCCTGAAGGATTACCCCTGCATGACCACTGCGGGAGGAGGCCACGGCCTTC	576	
Db	638	GAACAGATCCGACAGAGATTACCCCTGCACGCCACCTTCGCGGAAGAGGCCAGAGGTG	697	
QY	577	CTGCGTGGTCTGGGGTG-----AATGGAGCCAGCCGAGTACTTTGTGGGTGTCCAT	630	
Db	698	CTGGGTACGCTCGCCTGGGCGCGACAGGGAGCCGCCGCGACCTTTGTTCGGCGCTCAC	757	
QY	631	GTGCGCGGAGGGACTATGTGCATGTTCATGCCTTAATGTGTGGAAGGCGGTGGTGGCTGAC	690	

Db 758 GTGCGCGTGGGACTATCTCAGGTTATGCTCAGCGCIGGAAGGTTGTTGGGGAC 817
Qy 691 CGGGTTACCTGGAAGGCGCTGGATATGTTCCGGCGACGCTATCACTCCAGTCTTC 750
Db 818 AGCGCTACCTCCGGCAGGCGATGAGTGGTTCCGGCGACGCGACGAAGCCCGTTTC 877
Qy 751 GTGGTTACAAGCAACGGTATGGCTGGTGGCGGAGAACATTAATGCTTCCGAGGAGAC 810
Db 878 GTGTCACCAACAGCGATGAGTGGTTGAAGAAACATCGACCTCCAGGGCGAT 937
Qy 811 GTGTTGTCGGCGCAATGATGAGGGGTCGCCAGCAAGAGACTTGGCGTCTACCTACC 870
Db 938 GTGAGGTTGCTGGCGATGAGAGAGGCTACACCGTGGAAAGACTTTGGCCIGCTACA 997
Qy 871 CAGTGCACACACATCATCATGACTATGGGACCTTGGGATTTGGCTGGCTACCTGGCA 930
Db 998 CAGTGCACACACATCATCATGACTATGGGACCTTGGGCTTCTGGGCTGGCTACCTGGCT 1057
Qy 931 GTGGTGTATCACTACTAGCCAACTACACCTTCCGGATTCTCCGTTTCCICAAAGTC 990
Db 1058 GCGGAGACACTGCTACTTGGCCAACTTACACCTGCCAGACTCTGAGTTCTTGAAGATC 1117
Qy 991 TTTAAGCAGAGGAGAGCTTCTTACCGAATGGGTGGGATCCCTGCGGATCTGTCCCA 1050
Db 1118 TTTAAGCGGAGGCGCTTCTTCCCGAGTGGGTGGGCAATTAATGAGACTTGTCTCA 1177
Qy 1051 CTC 1053
Db 1178 CTC 1180

RESULT 13

US-09-863-475A-5
; Sequence 5, Application US/09863475A
; Patent No. US20020102688A1
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/863,475A
; FILING DATE: 24-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248955 OPAT JR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 8174 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-863-475A-5

Query Match 35.2%; Score 405; DB 10; Length 8174;
Best Local Similarity 68.4%; Pred. No. 6.2e-119;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;
Qy 217 GGCATGTTTCAGATCAATTCATTCGCGCGCTGGGAAACCGATGGCGCAATACCCACA 276
Db 4920 GGCACCTGGACGTGTACCCCAATGGCGGTTTGGTAATCAGATGGGACAGTATGCCACG 4979
Qy 277 CTCCTTTGACATGGCGAGGATGAACGAGCGGTTCGTTTCATCCCCCGCATCATGACAAAC 336
Db 4980 CTGCTGGCTCTGGCCAGCTCAACGCGCGCGGCGCTTTATCTCTGCTTGCATGCATGCC 5039
Qy 337 GCTCTAGCGCCCATTCCTCAGGATCAGCTCCCGGTGTTTACACAGCGACACGGCAAAAG 396
Db 5040 GCCCTGGCGCGGTATTCGGATCACCCTGCCCTGCTGGGCCCGCAGAAAGTGCACGCCG 5099
Qy 397 ATCCCATGGCAGAAATTACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATCCG 456
Db 5100 AGCGGTGGCGGAGCTGACGCTTACGACTGGATGCGGAGGATACCGGCTTTGAGA 5159
Qy 457 GGACACTTTGTGCGCTTACCGGGATACCCGTCTCTGAGCTTCTTACACACACTTCGCGC 516
Db 5160 GATCCTTTCCGAAAGCTCTCTGGCTTCCCTGCTTGGACTTCTTCCACCATCTCGG 5219
Qy 517 CCAGAGATCCGAGGAGTTCACTCAGCTGATGACACAGCTGCGGGAGGAGGCGCCGCTTC 576
Db 5220 GAACAGATCCGAGAGGTTCACTCAGCTGATGACACACTTCGCGGAGGAGGCGCAGAGTGT 5279
Qy 577 CTGGTGTGCTGCGGGTG-----AATGGAGCAGCGAGTACTTTTGTGGGTGTCCAT 630
Db 5280 CTGGGTGAGTCCCGCTGGCGGCGACAGGAGACCGCGCGACCTTTGTGCGGCTGCAC 5339
Qy 631 GTGCGCGAGGGGACTATGTCATGCTCATGCTTAATGTGTGAAGGGCGTGTGCTGCTGAC 690
Db 5340 GTGCGCGGTGGGACTATCTCAGGTTATGCTCAGCGCTGAGAAAGGTGTGTTGGGCGAC 5399
Qy 691 CGGGTTACCTGGAAGGCGCTGATGTTCCGGGACCGCTATCTATCTCCAGTCTTC 750
Db 5400 AGCGCTTACCTCCGCGAGGCGATGCTTCCGGGACCGGACGAGGCCCGCTTTC 5459
Qy 751 GTGGTTACAAGCAACGGTATGGCTGTGCGGAGAACATTAATGCTTCCGAGGAGAC 810
Db 5460 GTGTCACCAACAGCGATGAGTGTGTAAAGAAACATCGACCTCCAGGGCGAT 5519
Qy 811 GTGCTGTTCGGCGCAATGTTATGAGGGTTCGCCAGCAAGAGACTTCGCGCTCTCACC 870
Db 5520 GTGACGTTTGTGCGGATGGACAGAGGCTACACCGTGGAAAGACTTTGCCCTTCICACA 5579
Qy 871 CAGTGCACACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCTACCTGGCA 930
Db 5580 CAGTGCACACACACCATCATGACTATTTGGGACCTTTGGGCTTCTGGGCTGCTACCTGGCT 5639
Qy 931 GGTGTGATACCATCTACTTAGCCAACTACACCTTCCGGATCTCCGTTCTCTCAAGTC 990
Db 5640 GCGGAGACACTGCTACTTGGCGCACTTACCTTCCAGACTCTGAGTCTGATCTCTGAAGATC 5699
Qy 991 TTTAAGCAGAGGAGCGCTTCTTACCCGAAATGGGTGGGCAATCCCTGCGGATCTGTCCCA 1050
Db 5700 TTTAAGCGGAGGCGCTTCTTCCCGAGTGGGTGGGCAATTAATGAGACTTGTCTCCA 5759
Qy 1051 CTC 1053
Db 5760 CTC 5762

RESULT 14
US-09-999-672-29
; Sequence 29, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-09-999-672-29

Query Match 6.3%; Score 72.8; DB 10; Length 100;
Best Local Similarity 83.0%; Pred. No. 2.4e-13;
Matches 83; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 827 ATGGTATTGAGGGTCCGACCAAGACTTCGGCTGCTCACCAGTGTCAACCCACCA 886
Db 1 ATGGCTTCGAGAGCTCGCGGCCAAGACTTTGGCTGCTCAGCAGGTATACCCACCG 60
QY 887 TCATGACTATTGGGACCTTTGGGATTTGGGCTGCCTACCT 926
Db 61 TCATGACCATCGGCACCTTTGGGATCTGGCGCCCTACCT 100

RESULT 15
US-10-040-863-29
; Sequence 29, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-10-040-863-29

Query Match 6.3%; Score 72.8; DB 12; Length 100;
Best Local Similarity 83.0%; Pred. No. 2.4e-13;
Matches 83; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 827 ATGGTATTGAGGGTCCGACCAAGACTTCGGCTGCTCACCAGTGTCAACCCACCA 886
Db 1 ATGGCTTCGAGAGCTCGCGGCCAAGACTTTGGCTGCTCAGCAGGTATACCCACCG 60
QY 887 TCATGACTATTGGGACCTTTGGGATTTGGGCTGCCTACCT 926
Db 61 TCATGACCATCGGCACCTTTGGGATCTGGCGCCCTACCT 100

Search completed: May 26, 2003, 15:19:53
Job time : 178.211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:09:53 ; Search time 1886.5 Seconds

(without alignments)
9864.119 Million cell updates/sec

Title: US-10-040-863-7

Perfect score: 1149

Sequence: 1 atggcagcgccaggttcc.....gaagagcctctgatggaaa 1149

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_Other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.0	35.7	656	10	BB624949
2	398.6	34.7	1133	14	BQ920588
3	343.8	29.9	714	10	BE266792
4	283	24.6	596	14	BM744084
5	275.4	24.0	590	14	BM764661
6	273.2	23.8	591	14	BM781532

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	272.4	23.7	587	14	BM764788
8	271.8	23.7	784	9	AL046138
9	265.4	23.1	552	14	BM764751
10	247	21.5	521	12	BF710517
11	243	21.1	440	10	BB849545
12	232.2	20.2	449	17	BB274506
13	217	18.9	522	14	BM773356
14	212.2	18.5	895	10	BE299719
15	208.4	18.1	436	10	BB848859
16	207.8	18.1	541	12	BF194101
17	190.2	16.6	513	12	BE756417
18	187	16.3	326	10	AW843917
19	186.4	16.2	463	12	BF358052
20	176.8	15.4	828	12	BG192820
21	167.4	14.6	537	12	BE750891
22	159.6	13.9	207	10	AW347661
23	150.6	13.1	596	13	BJ091033
24	136.4	11.9	578	17	AZ285162
25	136	11.8	383	12	BF930329
26	122.4	10.7	424	13	BI534769
27	120.4	10.5	655	13	BJ040991
28	107.2	9.3	837	12	BF304028
29	106.8	9.3	521	17	BF342474
30	99.4	8.7	366	10	BE667589
31	96.8	8.4	526	10	BE019528
32	80	7.0	279	10	BB850127
33	75.8	6.6	649	10	AW917919
34	75.6	6.6	224	10	AW838837
35	75.6	6.6	484	13	BI350052
36	72.6	6.3	518	17	BF342476
37	67.6	5.9	264	14	BM742943
38	67.6	5.9	300	9	AU100082
39	67.6	5.9	304	9	AL709188
40	67.6	5.9	728	14	BM974044
41	65.6	5.7	554	12	BG610560
42	64.4	5.6	265	14	BM758349
43	57.4	5.0	541	12	BF418108
44	51.2	4.5	475	9	AI606683
45	50.4	4.4	578	9	AA108903

ALIGNMENTS

RESULT 1	BB624949	656 bp	mRNA	linear	EST 31-AUG-2001
LOCUS	BB624949	RIKEN full-length enriched, adult male colon	Mus musculus		
DEFINITION	CDNA clone 90304201.11 5', mRNA sequence.				
ACCESSION	BB624949				
VERSION	BB624949.1	GI:15398432			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 656)				
	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.				
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216				

Email: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y., and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Yamanaka,I., Kiyosawa,H., Kondo,S., Saigo,T., Shinagawa,A., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
 Ishii,Y. and Hayashizaki,Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct.
 Func. Genomics* 2 pre, 1172-1186 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

es.
tion/Qualifiers
556
ism="Mus musculus"
lin="C57BL/6J"
ref="taxon:10090"
e="9030420L11"
e_lib="RKEN full-l
"male"
ue_type="colon"
stage="adult"
host="DH10B"

```

Note—Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGGATCTCTTTTTCCTTTTTCVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGATTAAATTAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phagescript KS(+) after bulk excision from Lambda PTC I. Cloning sites, 5' end: SalI 3' end: BamHI.

BASE COUNT		BASE COUNT		BASE COUNT	
ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN
156 a	204 c	156 g	138 t	2 others	

Query Match	35.7%;	Score 410;	DB 10;	Length 656;
Best Local Similarity	86.1%;	Pred. No. 1.7e-102.		

[illegible]

QY 61 GTGACTTCCACCATCATCCACCCTCCAGCAGGAATAGTGAGCTCCACCCCTGTGAG 120

Db 175 GTGACTCCACCATCATCCACCTCCAGAACGAATAGTGAAGCTCCAAACCGTGTGAGAG 234

121 AAGGAATTACCGATGACGACTCAAAATGTCCTCGGNAACACAGAAAGCCAGAGATCGGA 180

235	Db	AGGAATTAC---AGCGGTCAAATGTCTCACCAAAACCGCAAGACAGACATG---288
181	QY	CGGACAGCAGCAGCATGGGAATGGAGAGCTGGGGGCAATGTTACAGATCAATTCATT240
289	Db	-----CAGCAGAGTGCCAGGACTCGAGGGCATATTCAGGATCAATTCATC333
241	QY	GGCGGCGTGGGAACACAGATGGCGGAATACGCCACACACTCTTTCACATGGCCAGGATGAAC300
334	Db	GGCGGCTGGGAACACAGATGGCGGAATATGCTACATTTGTTCACTGGCCAGGATGAAC393
301	QY	GGAGGCTTCGCTTCATCCCGCATCATGCACAACGCTCTACGGCCCATCTTCAGGATC360
394	Db	GGTGGGTTGCCCTTCATCCCTGAATCATGACACAGCTCTACGGCCCATCTTCAGGATC453
361	QY	AGCTCCCGGTGTATACACAGCGACACGGCCAAAAGATCCCATGGCAGAAATACCATCTC420
454	Db	AGTCTCCGGTGTTACACAGCGACACAGCCAGAGGATCCCGTGGCAGAAATACCACTTC513
421	QY	AAGCACTGGATGAGGAGCGGTTACCGCCACATTCGGGGACACTTGTGGCGTTCACGGGA480
514	Db	AAGCACTGGATGAGGAGCGGTTACCGCCACATCCCGGGCAGTATGTGGCTTTCAGGGGA573
481	QY	TACCCGTGCTCTGGGACCTTCTACCAACCACTCGCGCCACAGATCCTGAAGGAGTTCACC540
574	Db	TACCCGTGCTCTGGACCTTTTACCACCACTGGCCCAAGAAATCTGAGGAGTTCACC633
541	QY	CTGCATGACCAACGTGGCGGAGG562
634	Db	CTGCAGGACCATGTGGTGAGG655

RESULT 2
BQ920588/c

LOCUS	BQ920588	1133 bp	mRNA	linear	EST 20-AUG-2002
DEFINITION	AGENCOURT_10016100	NIH_MGC_101	Homo sapiens	cDNA clone	
IMAGE:6497629	5'		mRNA sequence.		

ACCESSION BQ920588
VERSION BQ920588.1 GI:22335286

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Mammalia; Eutheria; Primates; Catarrhini

REFERENCE
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
I (bases 1 to 1133)

TITLE National Institutes of Health, Mammalian Gene
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Straussberg, Ph.D.
Email: graphs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

recp://amagc:11m1:90v
 Plate: LLCM2678 row: p column: 14

FEATURES	Location/Qualifiers
High quality sequence stop:	

```
source
1. .l133
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:6407620"
```

```

/clone= IMAGE:0497029
/clone_lib="NIH_MGC_101"

```

```
/tissue_type="epidermoid carcinoma, cel
/lab_host="DH10B (phage-resistant)"
```

```
/note="Organ: lung; Vector: pOTB7; site_1: EcoRI; site_2: XhoI; cDNA made by oligo-dT priming. Direction: 5' to 3'."/>

```

into EcoRI/XhoI sites using the following 5' adapt

biochemist(s): Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California)

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a

```
BASE COUNT      268 a   332 c   312 g   221 t
ORIGIN
Query Match      34.7%; Score 398.6; DB 14; Length 1133;
Best Local Similarity 80.6%; Pred. No. 3.3e-99;
Matches 478; Conservative 0; Mismatches 114; Indels 1; Gaps 1;
QY 333 CAACGCTTAGCGCCCATCTTCAGAT-CAGCCTCCGGTGTGTACACAGCCAGCGCCA 391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
593 CAGCACCCTGGGCCCATCTTCAGATCAACCTTCGCCGTGTGTGAAGCGCCAGCTCCA 534
QY 392 AAAAGATCCATGGCAGAAATACCATCTCAACGACTGATGAGGAGGCTTACC GCCACA 451
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
533 GCAGGATCCCTGGCAGAACTACCACTGAGCTGAGTCTGATGGAGGAGNAATACGCCACA 474
QY 452 TTCGGGACACTTTGTGGCTTCACGGGATACCCGTCTTCGCGGACCTTCTACCAACAC 511
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
473 TCCCGGGGAGTACGTCGGCTTCACCGGTACCCCTGCTCTAGACCTTCTACCAACACC 414
QY 512 TCGGCCACAGATCTGAAGGAGTTACCCCTGCATGACACGTCGCGGAGGAGGCCAGG 571
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
413 TCCGCCAGAGATCTTCAGGAGTTCACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
QY 572 CTTCTCTGCTGCTGGGCTGAATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 AGTTCTCGGGGCTGCAGGTGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 294
QY 632 TCCGCCAGGAGGACTATGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
293 TFCGCCAGGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 234
QY 692 GGGGTTACTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
233 GCGGATACCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 174
QY 752 TGGTTACAGCAACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 TGGTCACCAAGTAAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 114
QY 812 TGGTGTTCGGGGCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 871
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 TGGTGTTCGGGGCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 54
QY 872 AGTGAACACACACCACTATGACTATTGGGAGCTTTGGGATTTGGGCTGCTGCTGCT 924
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 AGTGAACACACACCACTATGACTATTGGGAGCTTTGGGATTTGGGCTGCTGCTGCT 1

/clone="IMAGE:3534501"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      146 a   233 c   203 g   132 t
ORIGIN
Query Match      29.9%; Score 343.8; DB 10; Length 714;
Best Local Similarity 73.2%; Pred. No. 3.9e-84;
Matches 497; Conservative 0; Mismatches 147; Indels 35; Gaps 3;
QY 9 CCGCCAGGTTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 66
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 CCGTCAGATGCCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 111
QY 69 CACCATATCCACCTCCAGCAGCAATAGTGAAGCTCAACACCTCTGTACAGAAAGGAAT 128
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112 CACTATATTTACGTTACGACGCGGCTAGCAAGATTCAGCCATGTGGAGGTACCGGT 171
QY 129 ACCGATACGACTCAATGTCTCTCGGGAACACAGAAAGCCAGAGATCGCAGGAGAC 188
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 GCAGATACCATGCTAGCTCAACATCAAGCACTGGGACCCA----- 215
QY 189 CGAGCAGCATGGGATGAGAGCTGGGGGCACTTTCAGCATCAATTCATTTGGCGGCT 248
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 -----GCCAGCTCAGGGGATGTGACGATCAATTCAGCGGCT 258
QY 249 GGGGAACAGATGGGCAATAGCCACACTTTTGACCTGGCCAGGATGAACGAGCGGT 308
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 GGGGAACAGATGGGCAATAGCCACACTTATGCTTGGCCAGGATGAACGAGCGGCT 318
QY 309 TCGCTTATCTCCCGCATCCATGCAACACCTTAGCGCCCATCTTTCAGATCAGCTCC 368
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
319 CGCTTATCTCCCGCATCCATGCAACACCTTAGCGCCCATCTTTCAGATCAGCTCC 378
QY 369 GGTGTACACAGGACAGCGCCCAAGATCCATGCGAGATTAACCATCTCAACAGCTG 428
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
379 GGTGTACACAGGACAGCGCCCAAGATCCATGCGAGATTAACCATCTCAACAGCTG 438
QY 429 GATGAGGAGCGTTTACCGCCACATTCGGGCACTTTGCGCTTCACGGGATACCGGTG 488
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 GATGAGGAGCGTTTACCGCCACATTCGGGCACTTTGCGCTTCACGGGATACCGGTG 498
QY 489 CTCCTGGACCTTCTTACCAACCTCGGCCAGAGATCTTGAAGGAGTTTACCTTCGATGA 548
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
499 CTCCTAGACCTTCTTACCAACCTCGGCCAGAGATCTTCCAGGAGTTTACCTTCGATGA 558
QY 549 CCACCTGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 508
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
559 CCACCTGCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 617
QY 609 GAGTACTTTTGGGTGTCTCATGTGCGGCGGA-GGGGACTATGTGATGTCATGCTCAATG 667
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618 GGGCACCCTTTGAGGGTCCATGTTGCGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGG 677
QY 668 TGTGAAGGCGGTGGTGGC 686
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678 TGTGAAGGCGGTGGTGGC 696

RESULT 4
BM744084
LOCUS
DEFINITION K-EST0017531 S3SNU16 Homo sapiens cDNA clone S3SNU16-14-A06 5',
linear
596 bp mRNA
EST 01-MAR-2002
```


ECOR1 which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drf)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT	123 a	201 c	153 g	113 t
ORIGIN				
Query Match	24.08;	Score 275.4;	DB 14;	Length 590;
Best Local Similarity	71.24;	Pred. No. 3.1e-65;		
Matches	394;	Conservative	0;	Mismatches 126; Indels 33; Gaps 1;
Qy	9	CGCCAGGTTCCCTTCCTTCCTCTGCGCCACTTCCTCATCTTCTTCGTGACTTC	68	
Db	71	CGTTCAGATCCCTTCCTTCCTTCCTGCGCCACTTCCTTCCTTCGTGACTTC	130	
Qy	69	CACCATCATCCATCCATCCAGCAGGAATAGTGAAGCTCCCAACCCCTTCAGAGAGGAATT	128	
Db	131	CACATATATTCACGTTTCAGCAGGGGCTAGCGAAGATTCAGCCATGTGGGAGTTACCGGT	190	
Qy	129	ACCATCAGCACTCAATGTCCTCGGGAACACAGAAAGCCAGAGATGCGACGGGACAG	188	
Db	191	GCAGATACCACTAGTCTACCCCTCAACATCAAAGGCACCTGGGACCCA	234	
Qy	189	CGAGCAGCATGGGAATGAGAGCTGCGGGGCATGTTTCACGATCAATTCATTTGCCGGCT	248	
Db	235	-----GCCAGCTCAGGGGATGTGGAGCATCAATCAATAGGCCGCT	277	
Qy	249	GGGAACCATGAGGGGGAATACGCCACACTCTTTGTCACTGGCCAGGATGAACGGACGGCT	308	
Db	278	GGGAACCATGAGGGGGAATACGCCACACTCTACGCCCTGGCCAAGATGAACGGGCGCC	337	
Qy	309	TGGTTTCATCCCGCATCCATGACACAGCTCTAGCCCATCTTCAGGATCAGCCCTCC	368	
Db	338	CGCTTCATCCCGCCAGATGACACAGCCCTGGCCCATCTTCAGATCAGCCCTGCC	397	
Qy	369	GGTGTTCACAGCGACACGGCCAAAAGATCCCATGGCAGAAATTAACCATCTCAACGACTG	428	
Db	398	GGTGTTCACAGCGCCACGGCCAGCAGATCCCTGGCAGAACTACACCTGAATGACTG	457	
Qy	429	GATGGAGAGCGTTACCGGCACATTCGGGACACTTTGTGGCTTCACGGGATACCCGCTG	488	
Db	458	GATGGAGAGGAATACCGGCACATTCGCGGGGAGTACGTCGCTTCACCGGCTACCCCTG	517	
Qy	489	CTCCTGACCTTCTACCAACACCTGCCCCAGAGATCCTGAAGGAGTTTCAACCTGCAATGA	548	
Db	518	CTCCTGACCTTCTACCAACACCTGCCCCAGAGATCCTTCAGGAGTTTCAACCTGCAATGA	577	
Qy	549	CCAGTCCGGGAG	561	
Db	578	CCAGTCCGGGAG	590	

RESULT 6
BM781532
LOCUS
DEFINITION K-EST0058109 S3SN16s1 Homo sapiens cDNA clone S3SN16s1-19-B07 5', mRNA sequence.
BM781532
ACCESSION

BM781532.1 GI:19129764
EST.
human.
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 591)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and K.M.Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 19 row: B column: 07
High quality sequence stop: 591.
Location/Qualifiers
1. 591
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S3SN16s1-19-B07"
/clone_lib="S3SN16s1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pTZ18Rpl; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analysing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drf)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT	123 a	201 c	150 g	117 t
ORIGIN				
Query Match	23.84;	Score 273.2;	DB 14;	Length 591;
Best Local Similarity	70.94;	Pred. No. 1.3e-64;		
Matches	393;	Conservative	0;	Mismatches 128; Indels 33; Gaps 1;
Qy	9	CGCCAGGTTCCCTTCCTTCCTTCCTGCGCCACTTCCTCATCTTCTTCGTGACTTC	68	
Db	71	CGTTCAGATCCCTTCCTTCCTTCCTGCGCCACTTCCTTCCTTCGTGACTTC	130	
Qy	69	CACCATCATCCATCCATCCAGCAGGAATAGTGAAGCTCCCAACCCCTTCAGAGAGGAATT	128	

[illegible]

RESULT 7
BM764788
LOCUS
DEFINITION
K-ES0046355 S3SNU16s1 Homo sapiens cDNA clone S3SNU16s1-4-D03 5',
587 bp mRNA linear EST 04-MAR-2002
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens
human.
BM764788
BM764788.1 GI:19094403
EST.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 587)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 4 row: D column: 03
High quality sequence stop: 587.
Location/Qualifiers
I. .587
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S3SNU16s1-4-D03"
/clone_lib="S3SNU16s1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"

```

/lab_host="Top10F"
/notes="Origin: Stomach; Vector: pTZ19RP1; Site.1: EcoRI;
Site.2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dt)14 as 3' primer. The PCR
products was used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F' with
electroporation method."

BASE COUNT      122 a 201 c 151 g 113 t
ORIGIN

Query Match      23.7%; Score 272.4; DB 14; Length 587;
Best Local Similarity 71.1%; Pred. No. 2.1e-64;
Matches 391; Conservative 0; Mismatches 126; Indels 33; Gaps 1;

Qy  9  CGCCAGGTCCTTTCTCTTTCTCTGCGCCACTCTCTCATCTTTGCTTCGTGACTTC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  71  CGTTCAGATGCCCTTCTCTCTTCCTCATGCGCCACTCTCTCTTTTACGGTTC 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  69  CACCATCATCCACTCCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAAGAGAAAT 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  131  CACTATATTTCAGTTCAGCAGCGGCTAGCGAGATTCAAGCCATGTGGAGTTACCGGT 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  129  ACGGATGACAGCTCAAAATGTCTCTCGGGAACACAGAAAGCCCGAGAGATGGACGGACAG 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  191  GCAGATACCAAGTGTAGCCCTCAACATCAAAAGGCACCTGGGACCA----- 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  189  CGAGCAGCATGGGAATGGAGAGCTGCGGGGCATGTTACAGATCAANTCCATTGCGCGGT 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  235  -----GCCAGCTCAGGGGGAGTGTGGACGATCAATGCATAGAGCGCGCT 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  249  GGGGAACAGATGGGGAATACGCCACACTCTTTGCACTGGCCAGGATGAACGAGCGGT 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  278  GGGGAACAGATGGGAGTAGCCCACTGTACGCCCTGGCCAGATGAACGGGCGGC 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  309  TGGGTTCAATCCCGCATCCATGCACACAGCTGTAGCGGCCATCTCAGGATCAGCCTCC 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  338  CGCCTTCAATCCCGGCCAGATGCACAGCAAGCCCTGGGCCCATCTTCAGAAATCACCCGTGC 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  369  GGTGTTACAGCGACAGCGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACAGCTG 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  398  GGTGCTGCAGCGCCACCGCCAGCAGGATCCCTGGCAGACTACCACTGATGACTG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  429  GATGGAGGAGCGTTACCGCCACATTCGGSACACTTTGTCCGCTTACCGGGATACCCGCTG 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  458  GATGGAGGAGGAATACCGCCACTTCCCGGGGAGTACGTCGGCTTCACCGGCTACCCGCTG 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  489  CTCTCGAGCTTCTACACAGACACCGCCAGAGATCTCTGAGGAGTTCAACCTGATGA 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  518  CTCTCGAGCTTCTACACACACTCCGCCAGGAGATCTCCAGGAGTTCAACCTGATGA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY      549 CCACGTGGCG 558
      |||||
Db      578 CCACGTGGCG 587

AL046138      784 bp mRNA linear EST 29-FEB-2000
DKFP434P0672_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFP434P0672 5', mRNA sequence.
AL046138
ACCESSION AL046138.2 GI:5936190
VERSION EST
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Poustka.A., Klein.M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
JOURNAL EST (Poustka, et al.)
COMMENT Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5434225.
Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Resequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No SL sequence available.
This clone (DKFP434P0672) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             Location/Qualifiers
     source
     1..784
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="DKFP434P0672"
     /clone_lib="434 (synonym: htes3)"
     /tissue_type="testis"
     /dev_stage="adult"
     /lab_host="DH10B"
     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT      142 a      234 c      248 g      157 t
ORIGIN
Query Match      23.7%; Score 271.8; DB 9; Length 784;
Best Local Similarity 73.6%; Pred. No. 3.6e-54;
Matches 360; Conservative 0; Mismatches 127; Indels 2; Gaps 1;

QY      555 GCGGAGGAGGCCAGCGCTCTCGTGGTGTCTGGGGTGAATGGAGCGACCGCGATAC 614
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 GCGCGAGGAGGCCAGAACTTCTCGGGGCCCTGCGAGGCCAAGTGGGAGCGACGCGCA 60

QY      615 TTGTGTGGGTGCCATGTGCGCGGAGGAGCATGTGCATGTCCATTAATGTGGAA 674
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 CTTCGTGGGGTCCACGTGCGCGGGGGGACATGTCCGTGTGCATGTCCGCGGTATGGAA 120

QY      675 GGCCTGTGTGGTGACCGGGTTACTCGAAGAGGCCCTCGGATATGTTCCGGGACGCTA 734
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 GGGGGTCTCGGGGAGCC--GGGTACCTGACGCGGGCCCTGGAGTGTTCGGGCGCTGCTG 178

QY      735 TTCATCTCCAGTCTTCGTGTACRAGCAACGGTATGGCTTGTGCGCGGAGACATTAA 794
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      179 CGCCCTCCGGGCTTTGTGGTCACCAACGATGACATGGGCTGTGGTCCGGGAGAGATCAA 238

QY      795 TGCCTCCCGAGGAGACGTGTGTTCTCGGGGCAATGTTATGAGGGGTGCGCCCAAGGA 854

```

```

Db      239 CAGCTCCCTTGGGACGTGGTGTTCGCTGCAATGCGCTCCAGGGCTCACCTGCCAAGGA 238
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      855 CTTTCGGCGTGCTCACCCAGTGCACACACACACCATATGACTATTGGACCTTTGGGATTG 914
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      299 CTTTCGCACCTGCTCACACAGTGCACACACACCATATCACCCTGGGACCTTCGGGGTCTG 358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      915 GGTGCTGCTACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTACACCTTCCGATTC 974
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      359 GGCCTGCTACCTCGCGGGGGGACACTGTGTACCTGGCCAACTTACCCIGCCCAACTC 418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      975 TCCGTTCTCTCAAAGTCTTTAAGCCAGGAGCGAGCTTCCTACCCGAATGGTGGCATCCC 1034
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      419 CCCTTTCAAGTGTGCTTTAGGCCGTAGAGCGGCTTCTCTCCAGAGTGGTGGCCTTGC 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      1035 TGCCGATCT 1043
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      479 GGCIGACCT 487
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
BM764751
LOCUS
DEFINITION K-EST0046310 S3SNU16s1 Homo sapiens cDNA clone S3SNU16s1-5-H06 5',
mRNA sequence.
ACCESSION BM764751
VERSION BM764751.1 GI:19094366
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 592)
AUTHORS Kim,M.S., Hahn,J., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
JOURNAL 21C Frontier Korean EST Project 2001
COMMENT Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 5 row: H column: 06
High quality sequence stop: 592.

FEATURES             Location/Qualifiers
     source
     1..592
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="S3SNU16s1-5-H06"
     /clone_lib="S3SNU16s1"
     /sex="F"
     /tissue_type="Ascites"
     /cell_type="Lymphoblast-like"
     /cell_line="SNU-16"
     /lab_host="top10f"
     /note="Organ: Stomach; Vector: pTZ19RP1; Site_1: EcoRI;
     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
     bacterial alkaline phosphatase (BAP) and then decapped
     with tabacco acid pyrophosphatase (TAP). The decapped
     intact mRNA was ligated with DNA-RNA linker including EcoR
     I site by treatment of T4 RNA ligase and the first strand
     cDNA was synthesized from oligo dt-selected mRNA by
     priming with dt-tailed vector. The dt-tailed vector was
     adjusted to have about 60nt. The cDNA vector was
     circularized with E. coli DNA ligase after digestion of
     EcoRI which site is also included in vector. An RNA strand
     converted to a DNA strand by Okayama-Berg method. The
     obtained cDNA vectors were used for transfection of the
     competent cells E. coli top10f, by electroporation method.
     The cDNA libraries constructed by this method are

```

full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transfection of the remaining DNA into competent cells *E. coli* Top10F⁺ with electroporation method."

```

BASE COUNT      123 a   201 c   155 g   113 t
ORIGIN
Query Match      23.1%; Score 265.4; DB 14; Length 592;
Best Local Similarity 71.2%; Pred. No. 1.9e-62;
Matches 395; Conservative 0; Mismatches 126; Indels 34; Gaps 2;

Qy 9 CGCCAGGTTCTTCCTTCTCTGCTGCGCCACTTCCICATCTTTGTCGTGACTTC 68
Db 71 GTTCAGATGCTTCTCTTCCATGCGCCACTTCTCTCTTTGTTCACGGTTTC 130
Qy 69 CACCATCATCCACTCCAGCAGCGAATAGTGAAGTCCCAACCCCTGTCCAGAGAGGAATT 128
Db 131 CACTATATTTCACTTCAGCAGCGGCTAGCGAAGATTCAGCCATGTGGGAGTTACCGGT 190
Qy 129 ACCGATCAGCACTCAATGTCTCTGGGAACACAGAAAGCCAGAGATGCGACGGGACAG 188
Db 191 GCAGATACCACTAGTCACTCAACATCAAAGGCATCTGGGACCCA----- 234
Qy 189 CGAGCAGCATGGGAATGGAGAGTGGCGGCATCTTCACGATCAATTCATTCGCGGCT 248
Db 235 -----GCCAGCTCAGGGGATGTGGACGATCAATGCAATAGGCCGCT 277
Qy 249 GGGGAACAGATGGCGGAATACGCCACTTTTGCACCTGSCCCAGGTAGACGGACGCT 308
Db 278 GGGGAACAGATGGCGGAGTAGCCACACTGTACGCCCTGSCCAAGATGAACGGCGGCC 337
Qy 309 TGGCTTATCCCGCATCCATGCACACGCTCTAGCGCCCATCTTTCAGGATCAACCTCC 368
Db 338 CGCCTTATCCCGGCCAGATGCACACGACCCCTGGCCCATCTTTCAGATCAACCTGCC 397
Qy 369 GGTCTTACACAGCAGCAGCGCAAAAGATCCCAATGCGAGAAATACCATCTCAACGACTG 428
Db 398 GGTCTGACAGCGCCACGCGCAGCAGATCCCTGCGAGNACTACCACTGATGACTG 457
Qy 429 GATGGAGAGGTTTACCGCCCATTC- GGACACTTTGTGCGCTTACCGGATACCCGT 487
Db 458 GATGGAGAGGAATACCGCCACTTCCCGGGGGAGTAGCTCCCGTTCACCGGCTACCCCT 517
Qy 488 GCTCTGACCTTCTACCAACCTGCGCCAGAGATCCTGAAGGAGTTACCCCTGCATG 547
Db 518 GCTCTGACCTTCTACCAACCTGCGCAGGAGATCTTCCAGGAGTTACCCCTGCACG 577
Qy 548 ACCAGTGGGGAGG 562
Db 578 ACCAGTGGGGAGG 592

```

```

RESULT 10
LOCUS      BF710517/c
DEFINITION BF710517 521 bp mRNA linear EST 02-JAN-2001
            MI-P-AV1-nre-a-12-0-UI.s1 MI-P-AV1 Sus scrofa cDNA clone
ACCESSION  BF710517
VERSION     BF710517.1 GI:12009994
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 521)
Bonaldi,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktugle@iastate.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized placenta library cDNA Library Preparation: M.B. Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

Location/Qualifiers
1..521
/organism="Sus scrofa"
/strain="crossbreed"
/db_xref="taxon:9823"
/clone="MI-P-AV1-nre-a-12-0-UI"
/clone_lib="MI-P-AV1"
/lab_host="DH10B (Life Technologies)"
/notes="vector: pMT3D-pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: EcoRI; the MI-P-AV1 library is ultimately derived from placenta tissue. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/. The procedure used to create this library has been previously described (Bonaldi , Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_LIB=MI-P-AV1
TAG_TISSUE=placenta
TAG_SEQ=ATTGG"

BASE COUNT 120 a 119 c 170 g 112 t
ORIGIN

Query Match 21.5%; Score 247; DB 12; Length 521;
Best Local Similarity 79.1%; Pred. No. 2.2e-57;
Matches 306; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

```

Qy 711 CCTGGATGTTCCGGCAGCCTATTCTATCCAGTCTTCTGGTTACAAGCAACGGTAT 770
Db 521 CCTGGATGTTCCGGGCTGCTACCGCTCCCGCTTTTGGTCTCCAGACAGCAT 462
Qy 771 GGCCTGTTGTCGGGAGAACATTAATGCTTCCCGAGGAGACGTGTTCGGGGCAATGG 830
Db 461 GGCCTGTTGTCGGGAACAATCANTGCTCCGCGGCGATGTGTTCGGGCAATGG 402
Qy 831 TATTAGGGGTGCGCCAGCAAGAGCTTCGGCTGCTCACCCAGTGCACACACCATCAT 890
Db 401 CAICGAGGGTCCCGCCGCAAGACTTCGGCTGCTCAACGAGTGAACCAACATGTCAT 342
Qy 891 GACTATTGGGACCTTTGGGATTTGGCTGCTACCTGTCAGGTGGTGCATACCATCTACTI 950
Db 341 GACCATTTGGCAGCTTCGGGATCTCGCCGCGCATACCTTGTGTTGGAGAGACCATCTACT 282
Qy 951 AGCCAACCTACACCCCTTCGGGATTTCCGTTCCCTCAAAAGTCTTTTAAGCCAGAGCAGCTT 1010
Db 281 GGCCAAATTACAGCTCCCGGACTCTCCCTTCCTCAAACTCTTTAAGCCGAGGAGCAGCTI 222
Qy 1011 CCTACCCGAATGGGTGGGCATTCCTGCGGATCTGTCCCACTCTTAAAGGATTAACACC 1070

```

	sex=mixed	(tissue_type=hippocampus, dev_stage=adult, sex=male)	(tissue_type=whole body, dev_stage=9 days embryo, sex=mixed)	(tissue_type=lung, dev_stage=13 days embryo, sex=mixed)
BASE COUNT	103 a	134 c	99 g	104 t

Query Match 21.1%; Score 243; DB 10; Length 440;
Best Local Similarity 83.5%; Pred. No. 2.5e-56;
Matches 308; Conservative 0; Mismatches 40; Indels 2

D**b** 88 ATGGCGAGTGCCCGAGTACCTTTCCTGCTTTCCTCTGGGCCACCTTCCCATCATCTTGTCTCTT 14

QY 61 GTGACTTCCACCATCATCCACCTCCAGCAGCGGAATAGTGAAGCTCCAACCCCTGTCAGAG 12

Db 148 GTGACTTCCACCATCATCCACCTCCAGCAACGAATAGTGAAGCTCCAAACCTGTTCAGAG 20

Q_Y 121 AAGGATTAACCGATGACGACACCAATGTCTCGGGAAACACAGAAACCCAGACATGCCG 18

Db : 208 AAGGAATTAC---AGGGGGTTCAATGTCCTACCCAAACGCGGCAAGACAGACAT----- 26

QY 181 CGGGACAGCGAGGAGCATGGGAATGGACAGCTCGGGGGCAIGTTACAGATCAATTCATT 24

D6 261 -----GCAGCAGAGTGCCAAAGCTGAGGGGCAATATTCAGGATCATTCATCCATC

[illegible]

36	GGAGGGCTTCCGCTTCAATCCCGCATCCGCTACGGGCGAATGCTCAGGATC
04	301
DD	507
35	GGGCCCCCTGGGGGAACAGATGGGGCGGATATGCTTACATTTGTTTGCAC
35	GGGCCCCCTGGGGGAACAGATGGGGCGGATATGCTTACATTTGTTTGCAC

[illegible]

0V 361 AGCCTCCCG 369

427 AGTCTCCGG 435
D**b**

RESULT 12
BH274506

LOCUS	BH274506	449 bp	DNA	linear	GSS 30-NOV-20
DEFINITION	CH230-23L10.TJ	CHORI-230 Segment 1	Rattus norvegicus	genomic clone	

Accession	CH230-23110, DNA sequence.
BH274505	

VERSION BH2/4506.I GI:1/186908
KEYWORDS GSS.
SUBJECT

SOURCE	ORGANISM
Norway rat.	<i>Rattus norvegicus</i>

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Pattonae

radius.
1 (bases 1 to 449)
REFERENCE
AUTHORS Zhao S Shetty T Shatsman S Tsegaye C Geer K Shvartsheva

A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., Long, P., and Fraser, C. M.

TITLE	JOURNAL
Rat BAC End Sequences from Library CHORI-230 EcoRI segment	Unpublished (1999)

other_gss: CH230-23L10.TV
Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200

Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library

Clones may be purchased from BACPAC Resources
availability, please contact Pieter de Jong (paejongemail.cno.cfr)

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 23 row: 1 column: 10
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..449
/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-23110"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pBARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

BASE COUNT 92 a 117 c 132 g 108 t
ORIGIN

Query Match 20.2%; Score 232.2; DB 17; Length 449;
Best Local Similarity 74.8%; Pred. No. 2.5e-53;
Matches 291; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 598 GGGAGCCAGCGAGTCTTTGGTGTCCATGTGCCCGGAGGGGACTATGTGCATGTC 657
DB 58 GCATCCGCTCTCGTACCTATGTGGGTGTTCACGTGCGTCGTGGAGACTTTGGAGGTG 117
QY 658 ATGCCTAATGTGTGAAGCGGTGTGTGCTGACCGGGGTACCTGGAAAGGCCCTGGAT 717
DB 118 ATGCCTAACCCTGGAGGAGTGTGGTGACCGAGCTTACCTCCAGAAAGCCATTGAC 177
QY 718 ATGTTCCGGGACCGCTATATCTCCAGTCTTCGTGGTTTACAAGCAACGGTATGCCCTGG 777
DB 178 TGGTTCGGGGCGGCGCACAAAGACCCATCTTTGTGTCCACGAGCAATGGCATGAGATGG 237
QY 778 TGCGGGGAGACATTATAGCTTCCGAGGACGTGTGTCGGGGCAATGGTATTGAG 837
DB 238 TGTGTGGAGACATGACATATCCACGGTGTGGTGTGCTTCGCTGGCAATGGAGATGAG 297
QY 838 GGGTGGCCAGCAAGGACTTCGCGTGTCCACCGAGTCCACCAACACCATCATGACTATT 897
DB 298 GGTACGCGGGGAGGAGCTTCGCACTGTCTACCCAGTGTATATCAACACCATCATGACTATT 357
QY 898 GGGACCTTTGGATTGGCTTGGCTTACCTACGAGGTGTGTATACATCTACTATAGCCAAC 957
DB 358 GGGACCTTTGGCTTCTGGGCTGCTACCTGCTGTGGGAGACACTGTTTACCTTGTCTAAC 417
QY 958 TACACCCCTCCGGATTCTCCGCTTCTCTCAA 986
DB 418 TTCACCCCTTCAGATTCAGAGTTCTGAA 446

RESULT 13
BM773356
LOCUS BM773356 522 bp mRNA linear ESI 04-MAR-2002
DEFINITION K-EST0057795 S3SNUL6s1 Homo sapiens cDNA clone S3SNUL6s1-17-H05 5',
RNA sequence.
ACCESSION BM773356
VERSION BM773356.1 GI:19102971
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
CONTACT Contact: Kim YS
COMMENT Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 17 row: H column: 05
High quality sequence stop: 522.
Location/Qualifiers
1..522

FEATURES

Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S3SNUL6s1-17-H05"
/clone_lib="S3SNUL6s1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pTZ18Rpl; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of 14 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 - 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including 17
promoter as 5' primer and N(drr)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F with
electroporation method."

BASE COUNT 109 a 174 c 136 g 103 t
ORIGIN

Query Match 18.9%; Score 217; DB 14; Length 522;
Best Local Similarity 68.5%; Pred. No. 4.4e-49;
Matches 332; Conservative 0; Mismatches 120; Indels 33; Gaps 1;
QY 9 GCGCCAGGTTCTCTTCTCTTCTCCCTCGCCACATCTCTCTCTTGTCTTCGTGACTTC 68
DB 71 CGTTCAGATGCTTCTCTTCTCTTCTCCATGGCCACATCTCTCTTGTCTTCGTGACTTC 130
QY 69 CACCATCATCCACCTCCAGCGCAATAGTGAAGCTCCAAACCCCTGTCCAGAGAGAAAT 128
DB 131 CACTATATTTCAAGTTCAGCGCGCTAGCGAAGATTCAAGCCATGTGGAGTTACCGGT 190
QY 129 ACCGATGACGACTCAATGTCTCGGAAACACAGAAAGCCAGAGATGCGAGGGACAG 188
DB 191 GCAGATACAGTGTAGCTCAACATCAAGGCACTGGGACCA----- 234
QY 189 CGAGCAGATGGGATGGAGACTCGGGGGATGTTTCAGATCAATTCATTGCGCGGCT 248
DB 235 -----GCCAGCTCAGGGGATGTGGAGCATCAATGCAATAGCGCGGCT 277
QY 249 GGGGAACAGATGGCGGAATACGCCACACTCTTTCAGTGGCCAGAGTAACAGCGAGCT 308
DB 278 GGGGAACAGATGGCGGAGTACGCCACACTGTACGCCCTGGCCAGATGACAGCGCGGCC 337
QY 309 TCGGTTTCATCCCGCATCCATGCACACAGCTCTAGCGCCCATCTTTCAGGATCAGCTCCC 368

```

338 CGCTTTATCCCGCCGAGATGACAGCACCCCTGGCCCCCATCTTCAGAAATCAACCTGCC 397
369 GTGTTTACACAGGACAGCCGCAAAAGATCCCATGGCGCAATATCCATCTCAACGACTG 428
398 GTGCTGCACAGGCGCCAGCCAGAGATCCCTGGCAGAACTACCACTGAATGACTG 457
429 GATGGAGGAGCTTACCGCCACATTCGCGGACACTTTGTGCGCTTACGGGATACCGGTG 488
458 GATGGAGGAGGAATACCGCCACTTCCCGGGGAGTAGCTCGCTTACCGGCTACCCGTC 517
489 CTCCT 493
518 CTCCT 522

RESULT 14
LOCUS BE299719
DEFINITION BE299719 895 bp mRNA linear EST 20-JUL-2000
CDNA clone NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960318 5',
mRNA sequence.
ACCESSION BE299719
VERSION BE299719.1 GI:9183467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 895)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM53 row: h column: 15
High quality sequence stop: 758.
FEATURES
Location/Qualifiers
1..895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2960318"
/clone_lib="NIH_MGC_17"
/tissue_type="rhodomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; Site:1: EcoRI;
Site:2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

331 CACAACGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGTGTGTACAGCAGCACGGCC 390
342 CACAACGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGTGTGTACAGCAGCACGGCC 411
391 AAAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGCTTACCGCAC 450
342 AGCAGGATCCCTGGCAGAACTACCACTCAACGACTGGATGGAGGAGGAGTACCGGCAC 601
451 ATTCGCGGACACI-TTGTGGCTTCAGGGAT-ACCGTGTCTCTCGACCT---TCFACC 505
602 ATCCCGGGGCGCTGATGCTCGGCTACCGGCTAACCCCTGCTCTCTCGACCTCTACACCA 661
506 ACCACCTGCGCCGACAGATCC-TGAAGAGTTTACCCCTGC-ATGACACAGTGGGGGAGGA 563
662 CCAACCTCCGCGAGGAGATCTTCCAGAGATTACCTTGCAACGACCACTGCGCGAGGA 721
564 GGCCGAGCCTTCTCGGTGGTCTGGGGGTGAATGGGAGCCAGCGAGTACTTTTGTGG 623
722 GGCGAAGAACTTCTCGGGGCTGCGAGGCGCAAGTGGCGAGGCGAGCGACCTTCGTGG 781
624 TGTCATGTCGCGGAGGAGGACTATGTGCTGT 656
782 GTTCC-CGTGCGCGGGGAGTATGTCCGGGT 813

RESULT 15
LOCUS BB848859
DEFINITION BB848859 436 bp mRNA linear EST 26-NOV-2001
CDNA clone F930004014 5', mRNA sequence.
ACCESSION BB848859
VERSION BB848859
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 436)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hirose,T., Hirose,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Kondo,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okidc,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Wachiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9226
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Wachiki,M., Iweda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
```

encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES
source

Location/Qualifiers

```
1..436
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930004C14"
/clone_lib="RIKEN full-length enriched, adult inner ear"
/tissue_type="inner ear"
/dev_stage="adult"
/note="pooled tissues : (tissue_type=cerebellum,
dev_stage=16 days neonate, sex=mixed),
(tissue_type=cerebellum, dev_stage=0 day neonate,
sex=mixed), (tissue_type=hippocampus, dev_stage=adult,
sex=male), (tissue_type=whole body, dev_stage=9 days
embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days
embryo, sex=mixed)"
```

BASE COUNT 107 a 137 c 98 g 94 t

ORIGIN

```
Query Match 18.1%; Score 208.4; DB 10; Length 436;
Best Local Similarity 81.9%; Pred. No. 9.7e-47;
Matches 285; Conservative 0; Mismatches 41; Indels 22; Gaps 3;

QY 1 ATGGCCAGCGCCAGGTCCTTCCTTCCTCTGCGCCACCTTCCTCATCTTTGCTTC 50
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GTGACTTCACCATCATCCACTCCAGCAGGGAATAGTGAAGTCCAAACCCCTGTCCAG 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 171 GTCACCTCCACCATCATCCACTCCAGCAGGGAATAGTGAAGTCCAAACCCCTGTCCAG 229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 AAGGAATACCATGACGACTCAATGTCTCTGGGAAACACAGAAAGCCAGAGATGCGA 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 230 AAGGAATAC---AGGCGGTTCAATGTCTCTCCAAACCGGCGGCAAGACACAT--- 282
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CGGGACAGGAGCAGCATGGGAATGAGAGCTGCGGGGCATGTTCCAGATCAATTCAT 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GCGCGGCTGGGACACAGATGGGGAATACGCCACACTTTTGCACTGGCCAGGATGAAC 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 329 GCGCGGCTGGGACACAGATGGGGAATATGCTACATGTTTGCACTGGCCAGGATGAAC 388
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GGACGGCTTGCTTCATCCCGCATCCATGACACACGCTCTAGCGGCC 348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 389 GGTGGGCTTGCTTCATCCCTGAATCCATGACACACGCTCTAGCGGCC 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: May 26, 2003, 13:09:21
Job time : 1892.5 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: May 27, 2003, 14:52:27 ; Search time 48.1951 Seconds
(without alignments)
1050.630 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVFFSPLAHLFIVF.....RSHFLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2032	100.0	380	21 AAB36104	Rat hepatoma H35 c
2	2032	100.0	380	23 AAE16622	Rat hepatoma H35 c
3	1896	93.3	353	21 AAB36105	Rat hepatoma H35 c
4	1896	93.3	353	23 AAE16623	Rat hepatoma H35 c
5	1502	73.9	340	19 AAW37855	Porcine secretor t
6	1502	73.9	340	19 AAW53101	Pig secretor. Sus
7	1485.5	73.1	343	19 AAW69332	Human Sec2 protein
8	1475	72.6	344	21 AAB36106	Human Sec2 catalyt
9	1057.5	52.0	365	12 AAB37551	GDP-Fuc:beta-D-gal
10	1057.5	52.0	365	16 AAR80154	GDP-L-fucose-beta-

11	1057.5	52.0	365	16 AAR70422	2-Alpha-fucosyltra
12	1057.5	52.0	365	17 AAR70421	2-Alpha-fucosyltra
13	1057.5	52.0	365	17 AAR90572	Human H-transferr
14	1057.5	52.0	385	18 AAW23805	Human alpha 1,2 fu
15	1057.5	52.0	365	18 AAW13640	Human H-transferr
16	1057.5	52.0	365	21 AAY97279	Human alpha(1,2)-f
17	1049.5	51.6	365	15 AAR45936	A glycosyltransfer
18	1046.5	51.5	365	15 AAW53102	Pig H transferase.
19	1046.5	51.5	365	21 AAY79302	Pig alpha-1-2 fucc
20	1046.5	51.5	385	23 AAB47995	Swine alpha (1,2)
21	1044.5	51.4	385	20 AAR97356	Swine alpha (1,2) f
22	1044.5	51.4	385	20 AAW30630	Swine alpha-1,2-fu
23	995	49.0	357	20 AAY17969	X. laevis alpha-1.
24	328.5	16.2	110	21 AAG03530	Human secreted pro
25	141	6.9	300	21 AAY92213	H. pylori UA802 al
26	141	6.9	300	22 AAG64071	Helicobacter pylor
27	139.5	6.9	287	23 AAW51992	Bacteroides fragil
28	103	5.1	661	22 AAW39722	Propionibacterium
29	101	5.0	690	23 AAE20142	Human protein cont
30	97	4.8	317	13 AAR26061	Growth Factor Rece
31	96.5	4.7	355	21 AAY70408	C. elegans alpha-1
32	95	4.7	1515	22 ABL10624	Novel human diagno
33	93.5	4.6	402	22 ABB63790	Drosophila melanog
34	93	4.5	1073	22 ABG28599	Novel human diagno
35	92	4.5	984	22 AAG93762	Human polypeptide.
36	90.5	4.5	659	22 AAB46426	Bilin binding prot
37	90.5	4.5	2008	23 AAU11814	Cancer and neuroge
38	90.5	4.5	2306	23 AAU11817	Cancer and neuroge
39	89.5	4.5	2352	23 AAU11816	Cancer and neuroge
40	89.5	4.4	586	21 AAY54126	Amino acid sequenc
41	89.5	4.4	604	22 AAU46700	Propionibacterium
42	89	4.4	353	16 AAB61993	Porphyromonas geng
43	89	4.4	334	22 AAB93020	Human protein sequ
44	89	4.4	1031	22 AAU17082	Novel signal trans
45	89	4.4	2057	22 AAE11890	Angiogenesis assoc

ALIGNMENTS

RESULT 1
AAB36104
ID AAB36104 standard; Protein; 380 AA.
XX
AC AAB36104;
XX
DT 19-FEB-2001 (first entry)
XX
DE Rat hepatoma H35 cell alphasyltransferase.
XX
DE Rat; alphasyltransferase; cytotstatic; neuroprotective;
KW nontropic; gene therapy; Fucalpal-2galbetail-3galNac; immunotherapy;
KW immunosuppression; cancer; neurological disease;
KW Small cell lung carcinoma.
XX
XX Rattus norvegicus.
XX
XX WO200064464-A1.
XX
PD 02-NOV-2000.
XX
PF 23-APR-1999; 99WO-US07384.
XX
PR 23-APR-1999; 99WO-US07384.
XX
PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.
XX
PI Holmes EH, Sherwood AL;
XX
XX WPI: 2000-687262/57.
DR N-PSDB; AAC67965.
XX
XX New rat ganglioside GM1-specific alphasyltransferase, useful for

PT Preparation of fucosyl GM1 which is useful as a nutritional composition
 XX or immunotherapeutic for cancer and neurological diseases -
 PS Claim 1; Fig 5; 91pp; English.

CC The present sequence was given in a specification relating to an isolated
 CC rat ganglioside GM₁-specific alpha-2-fucosyltransferase protein. The
 CC protein or its cellular fraction is useful for synthesis of a molecule
 CC comprising Fucalpal-2Galbeta1-3GalNAc, a glycolipid, glycoprotein,
 CC glycolipoprotein or a free oligosaccharide comprising
 CC Fucalpal-2Galbeta1-3GalNAc. The method involves contacting
 CC alpha-2-fucosyltransferase with GDP-fucose and a molecule or glycolipid,
 CC glycoprotein, glycolipoprotein or oligosaccharide having a terminal
 CC Galbeta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM₁
 CC by contacting the protein with GDP-fucose and ganglioside GM₁. The
 CC obtained glycoproteins, glycolipoproteins, glycolipids and
 CC oligosaccharides are useful as nutritional compositions and fucosyl-GM₁
 CC is useful for inducing an immunotherapeutic or immunosuppressive action
 CC against cancer, neurological disease or small cell lung carcinoma.

XX Sequence 380 AA;

Query Match 100.0%; Score 2032; DB 21; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.7e-212;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAQVPFSPFLAHFLIFVFTSTIIHLQORIVKLPSEKELPMTQMSSGNTSPENR 60
 DB 1 MASAQVPFSPFLAHFLIFVFTSTIIHLQORIVKLPSEKELPMTQMSSGNTSPENR 60

QY 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120
 DB 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120

QY 121 SLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHFVRETCYPCSWTFYHHLRPEILKEFT 180
 DB 121 SLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHFVRETCYPCSWTFYHHLRPEILKEFT 180

QY 181 LHDHVREEAQAFRLGRVNGSQPSTFVGHVHRRGDYVHVMPNVKGVADRGYLEKALDM 240
 DB 181 LHDHVREEAQAFRLGRVNGSQPSTFVGHVHRRGDYVHVMPNVKGVADRGYLEKALDM 240

QY 241 FRARYSSPVFVVTSSNGMAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNIHTITG 300
 DB 241 FRARYSSPVFVVTSSNGMAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNIHTITG 300

QY 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVKFPEAAFLPEWVGIPADLSPLLKALTPACP 360
 DB 301 TFGIWAAYLAGDITIYLANYTLPDSPFLKVKFPEAAFLPEWVGIPADLSPLLKALTPACP 360

QY 361 RSHFHLKAGVTCYVAGRAF 380
 DB 361 RSHFHLKAGVTCYVAGRAF 380

RESULT 2
 ID AAE16622 standard; Protein; 380 AA.
 XX AAE16622;
 XX AAE16622;
 DT 09-APR-2002 (first entry)
 XX Rat hepatoma H35 cell alpha-2Fuct.
 XX Rat; alpha-2-fucosyltransferase; alpha-2Fuct; antisense therapy;
 KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
 KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM₁; cancer;
 KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM₁;
 XX cell transformation.
 OS Rattus norvegicus.

FH Key Location/Qualifiers
 FT Domain 1..27
 FT /note= "Intracellular/Transmembrane domain"
 FT Modified-site 199
 FT /note= "N-glycosylated"
 FT Region 221..380
 FT /note= "Region which overlaps rat FTB"
 FT Modified-site 265
 FT /note= "N-glycosylated"
 FT Modified-site 293
 FT /note= "N-glycosylated"
 XX US6329170-B1.
 PN 11-DEC-2001.
 XX 23-APR-1999; 99US-0298886.
 XX 23-APR-1999; 99US-0298886.
 XX (NWHO-) NORTHWEST HOSPITAL.
 XX Holmes EH, Sherwood AL;
 XX WPI; 2002-121132/16.
 XX N-PSDB; AAD27207.
 XX Rat hepatoma H35 cell alpha-2-fucosyltransferase, useful for producing
 PT GM₁-specific alpha-2-fucosyltransferase enzyme by recombinant
 FT techniques and for detecting oncogenic transformation of test tissues -
 XX Claim 3; Fig 5; 41pp; English.

CC The invention relates to rat GM₁-specific alpha-2-fucosyltransferase
 CC (alpha-2Fuct) enzyme and its corresponding nucleic acid. This nucleic
 CC acid is specific for a carbohydrate moiety found in ganglioside GM₁,
 CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc);
 CC saccharide. Alpha-2Fuct DNA is useful for producing rat alpha-2Fuct
 CC protein by recombinant techniques. Alpha-2Fuct DNA is useful for the
 CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
 CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM₁.
 CC Alpha-2Fuct DNA is useful for detecting oncogenic transformation which
 CC involves assaying for changes in expression of alpha-2 Fuct. Since
 CC alpha-2Fuct is activated in cell transformation, antisense sequences
 CC derived from alpha-2Fuct DNA are useful for inhibiting, suppressing
 CC or treating cancer. Alpha-2Fuct DNA is useful in gene therapy and
 CC antisense therapy. The present sequence is rat hepatoma H35 cell
 CC alpha-2Fuct.
 XX Sequence 380 AA;

Query Match 100.0%; Score 2032; DB 23; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.7e-212;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAQVPFSPFLAHFLIFVFTSTIIHLQORIVKLPSEKELPMTQMSSGNTSPENR 60
 DB 1 MASAQVPFSPFLAHFLIFVFTSTIIHLQORIVKLPSEKELPMTQMSSGNTSPENR 60

QY 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120
 DB 61 RSEQHGNGELRGMTINSIGRLGNMGGEYATLFALARMNGRLAFTPASMHNALAPIFRI 120

QY 121 SLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHFVRETCYPCSWTFYHHLRPEILKEFT 180
 DB 121 SLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHFVRETCYPCSWTFYHHLRPEILKEFT 180

QY 181 LHDHVREEAQAFRLGRVNGSQPSTFVGHVHRRGDYVHVMPNVKGVADRGYLEKALDM 240
 DB 181 LHDHVREEAQAFRLGRVNGSQPSTFVGHVHRRGDYVHVMPNVKGVADRGYLEKALDM 240

QY 241 FRARYSSPVFVVTSSNGMAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNIHTITG 300
 DB 241 FRARYSSPVFVVTSSNGMAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNIHTITG 300

Db 241 FRARYSPVFTVTSNGMWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNNHTIMTIG 300
 QY 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVFKEPAFLPEWVGIPADLSPLLKALTPACP 360
 Db 301 TFGIWAAYLAGGDTIYLANITLPDPSFLKVFKEPAFLPEWVGIPADLSPLLKALTPACP 360
 QY 361 RSHFHLKAGVTCYVAGRAF 380
 Db 361 RSHFHLKAGVTCYVAGRAF 380

RESULT 3
 AAB36105
 ID AAB36105 standard; Protein; 353 AA.
 AC
 XX
 XX
 DT 19-FEB-2001 (first entry)
 XX
 DE Rat hepatoma H35 cell alphas-2fucosyltransferase catalytic domain.
 XX
 KW Rat; alphas-2fucosyltransferase; cytostatic; neuroprotective;
 KW nootropic; gene therapy; Fucalphan-2Galbetal-3GalNAC; immunotherapy;
 KW immunosuppression; cancer; neurological disease;
 KW small cell lung carcinoma.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200064464-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-US07384.
 XX
 PR 23-APR-1999; 99WO-US07384.
 XX
 PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.
 XX
 PI Holmes EH, Sherwood AL;
 XX
 DR WPI; 2000-687262/67.
 DR N-PSDB; AAC67966.
 XX
 PT New rat ganglioside GM1-specific alphas-2fucosyltransferase, useful for
 PT preparation of fucosyl GM1 which is useful as a nutritional composition
 PT or immunotherapeutic for cancer and neurological diseases -
 XX
 PS Claim 2; Fig 3A; 91pp; English.
 XX
 CC The present sequence is given in a specification relating to a rat
 CC ganglioside GM1-specific alphas-2fucosyltransferase protein. The protein
 CC or its cellular fraction is useful for synthesis of a molecule comprising
 CC Fucalphan-2Galbetal-3GalNAC, a glycolipid, glycoprotein, glycolipoprotein
 CC or a free oligosaccharide comprising Fucalphan-2Galbetal-3GalNAC.
 CC The method involves contacting alphas-2fucosyltransferase with GDP-fucose
 CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or
 CC oligosaccharide having a terminal Galbetal-3GalNAC group. It is also
 CC useful for synthesis of fucosyl-GM1 by contacting the protein with
 CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
 CC glycolipoproteins, glycolipids and oligosaccharides are useful as
 CC nutritional compositions and fucosyl-GM1 is useful for inducing an
 CC immunotherapeutic or immunosuppressive action against cancer,
 CC neurological disease or small cell lung carcinoma.
 XX
 SQ Sequence 353 AA;

Query Match 93.3%; Score 1896; DB 21; Length 353;
 Best Local Similarity 100.0%; Pred. No. 1e-197;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2B LQORIVKLOPLSEKELPMTOMKSSGNTSPSPERMRDSEOHNGELRGMTINSIGRLGNQM 87
 Db 1 LQORIVKLOPLSEKELPMTOMKSSGNTSPSPERMRDSEOHNGELRGMTINSIGRLGNQM 60

QY 88 GEYATLFAALRMNGRLAFIPASMHENALAFIRISLPVLHSDTAKKIPQWNYHLNDQMEER 147
 Db 61 GEYATLFAALRMNGRLAFIPASMHENALAFIRISLPVLHSDTAKKIPQWNYHLNDQMEER 120
 QY 148 YRHIPGHFVREFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 207
 Db 121 YRHIPGHFVREFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQPSTFV 180
 QY 208 GYHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFTVTSNGMWCRENINAS 267
 Db 181 GYHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFTVTSNGMWCRENINAS 240
 QY 268 RGDVVFAGNGIEGSPAKDFALLTQCNNHTIMTIGTFGIWAAYLAGGDTIYLANITLPDPSF 327
 Db 241 RGDVVFAGNGIEGSPAKDFALLTQCNNHTIMTIGTFGIWAAYLAGGDTIYLANITLPDPSF 300
 QY 328 LKVFKEPAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF 380
 Db 301 LKVFKEPAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAGVTCYVAGRAF 353

RESULT 4
 AAE16623
 ID AAE16623 standard; Protein; 353 AA.
 AC
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Rat hepatoma H35 cell alphas-2fucosyltransferase catalytic domain.
 XX
 KW Rat; alphas-2fucosyltransferase; alphas-2Fuct; antisense therapy;
 KW galactose betal-3N-acetylgalactosamine; Galbetal-3GalNAC; glycolipid;
 KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
 KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
 KW cell transformation; catalytic domain.
 XX
 OS Rattus norvegicus.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 172
 FT Region /note= "N-glycosylated"
 FT /note= "Region which overlaps rat FTB"
 FT Modified-site 238
 FT /note= "N-glycosylated"
 FT Modified-site 286
 FT /note= "N-glycosylated"
 FT Modified-site 292
 FT /note= "N-glycosylated"
 XX US6329170-B1.
 PD 11-DEC-2001.
 XX
 PF 23-APR-1999; 99US-0298886.
 XX
 PR 23-APR-1999; 99US-0298886.
 XX
 PA (NWHO-) NORTHWEST HOSPITAL.
 XX
 PI Holmes EH, Sherwood AL;
 XX
 DR WPI; 2002-121132/16.
 DR N-PSDB; AAD27208.
 XX
 PT Rat hepatoma H35 cell alphas-2fucosyltransferase, useful for producing
 PT GM1-specific alphas-2fucosyltransferase enzyme by recombinant
 PT techniques and for detecting oncogenic transformation of test tissues -
 XX
 PS Claim 4; Fig 3; 41pp; English.
 XX

CC The invention relates to rat GM1-specific alpha1-2fucosyltransferase
 CC (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic
 CC acid is specific for a carbohydrate moiety found in ganglioside GM1,
 CC a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc)
 CC saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct
 CC protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the
 CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
 CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.
 CC Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which
 CC involves assaying for changes in expression of alpha1-2 Fuct. Since
 CC alpha1-2Fuct is activated in cell transformation, antisense sequences
 CC derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing
 CC or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and
 CC antisense therapy. The present sequence is rat hepatoma H35 cell
 CC alpha1-2Fuct catalytic domain.

XX Sequence 353 AA;

Query Match 93.3%; Score 1896; DB 23; Length 353;
 Best Local Similarity 100.0%; Pred. No. 1e-197;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 28 LQORIVKLPSEKELPMTQSSGNTSPENRRDSEHGNGELRGMTTINSIGRLGNQM 87
 DB 1 LQORIVKLPSEKELPMTQSSGNTSPENRRDSEHGNGELRGMTTINSIGRLGNQM 60
 QY 88 GEYATLALARMNGRLAIPASMHNALAIPRISLPVLSHTAKKIPQNYHLNDWMEER 147
 DB 61 GEYATLALARMNGRLAIPASMHNALAIPRISLPVLSHTAKKIPQNYHLNDWMEER 120
 QY 148 YRHIPGHPVRTGYPCSTFVHLLRPETLKEFTLHDHYREAZAQLRLGVNGSPSTFV 207
 DB 121 YRHIPGHPVRTGYPCSTFVHLLRPETLKEFTLHDHYREAZAQLRLGVNGSPSTFV 180
 QY 208 GVHVRGSDYVHMPNVKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCENINAS 267
 DB 181 GVHVRGSDYVHMPNVKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCENINAS 240
 QY 268 RGVVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTIGVWVAYLAGSDTYYLANYLTPDSPF 327
 DB 241 RGVVVFAGNGIEGSPAKDFALLTQCNHTIMTIGTIGVWVAYLAGSDTYYLANYLTPDSPF 300
 QY 328 LKVKPEAAFLPEWVGIPADLSPILKALTACPRSHFLKAGVTCYVAGRAF 380
 DB 301 LKVKPEAAFLPEWVGIPADLSPILKALTACPRSHFLKAGVTCYVAGRAF 353

RESULT 5

AAW37855

DB AAW37855 standard; Protein; 340 AA.

XX AC AAW37855;

XX DT 28-AUG-1998 (first entry)

XX DE Porcine secretor transferase (FUT2).

XX KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;
 KW transgenic animal; xenotransplantation; organ transplant.

XX OS Sus scrofa.

XX FT Key Location/Qualifiers

FT Domain 1..4

FT FT /note= "N-terminal cytoplasmic tail"

FT Domain 5..26

FT FT /note= "transmembrane domain"

FT Domain 27..340

FT FT /note= "C-terminal domain"

FT Modified-site 185..187

FT FT /note= "Asn is N-glycosylated"

FT FT 251..253

FT FT /note= "Asn is N-glycosylated"

FT

FT Modified-site 279..281
 FT FT /note= "Asn is N-glycosylated"
 FT Modified-site 305..307
 FT FT /note= "Asn is N-glycosylated"
 XX WO9807837-A1.
 XX PD 26-FEB-1998.
 XX PF 22-AUG-1997; 97WO-AU00540.
 XX PR 23-AUG-1996; 96AU-0001823.
 XX PA (AUST-) AUSTIN RES INST.
 XX PI McKenzie IFC, Sandrin MS;
 XX DR WPI; 1998-169148/15.
 XX DR N-PSDB; AAV29003.
 PT Nucleic acid encoding glycosyltransferase able to compete with
 PT second such enzyme - particularly used to reduce expression of
 PT unwanted carbohydrate epitope(s) on tissues intended for
 PT transplantation
 XX Claim 6; Fig 1A-B; 40pp; English.
 XX This polypeptide comprises porcine secretor glycosyltransferase
 CC (SE or FUT2), a type II integral membrane protein has high affinity
 CC for type I and type III substrates. Its amino acid sequence was
 CC deduced from the nucleotide sequence of a genomic DNA clone (see
 CC AAV29003) isolated from a pig liver library on the basis of homology
 CC to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.
 CC Expression of a glycosyltransferase, such as porcine Se, in a
 CC tissue results in reduced expression of unwanted carbohydrate
 CC epitopes on the tissue, especially porcine heart, liver, kidney or
 CC pancreas, rendering it more suitable for transplantation, i.e. less
 CC immunogenic and of increased immunological acceptability. A
 CC claimed method of producing a cell from a donor species that is
 CC immunologically acceptable to a recipient species involves reducing
 CC levels of carbohydrate on the donor cell that causes it to be
 CC recognised as non-self by the recipient by expressing a nucleic
 CC acid for a glycosyltransferase such as porcine Se in the cell.

XX Sequence 340 AA;

Query Match 73.9%; Score 1502; DB 19; Length 340;
 Best Local Similarity 79.68; Pred. No. 9.1e-155;

Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPFSPFLAHFLIFVVTSTIIHLQQRIVKLQPLSEKELPMTTQMSSGNTSPENR 60
 DB 1 MLSMQASPFPTGPFILFVFTASTIFHLQQRWVKIQP--TWELQWVQT--TSP--- 52
 QY 61 RDSCHGNGELRGMTTINSIGRLGNQMGCEYATLALARMNGRLAIPASMHNALAIPRI 120
 DB 53 -----SSPQLKGMWTINAIGRLGNQMGCEYATLALARMNGRPAFIPPEMHSILAPRI 106
 QY 121 SLPLVLSDTAKKIPQNYHLNDWMEERYRHIPGHVFRGTGYPCSWTFVYHLLRPETLKEFT 180
 DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGCVRTGYPCSWTFVYHLLRPETLKEFT 166
 QY 181 LHDHVREREAQFLRGLRVNGSQPSTFVGHVVRGSDYVHMPNVKGVVADRGYLEKALDM 240
 DB 167 LHNHVREREAQDFLRGLRVNGSRPSTYVGVHVRGSDYVHMPNVKGVVADRGYLEKALDW 226
 QY 241 FRARYSPFVVTNGMAWCENINASRGDVFVAGNGIEGSPAKDFALLTQCNHTIMTIG 300
 DB 227 FRARYSPFVVTNGMAWCENINASRGDVFVAGNGIEGSPAKDFALLTQCNHTIMTIG 286
 QY 301 TFGIWAAYLAGSDTYYLANYLTPDSPFLKVKPEAAFLPEWVGIPADLSPILK 353
 DB 287 TFGIWAAYLAGGETIYLANYLTPDSPFLKVKPEAAFLPEWIGIEADLSPILK 339

```

RESULT 6
AAW53101
ID AAW53101 standard; Protein; 340 AA.
XX
AC AAW53101;
XX
DT 08-JUL-1998 (first entry)
XX
DE Pig secretor.
XX
KW Pig; secretor; chimeric; glycosyltransferase; gene therapy;
transplantation.
XX
OS Sus scrofa.
XX
PN WO9805768-A1.
XX
PD 12-FEB-1998.
XX
PF 01-AUG-1997; 97WO-AU00492.
XX
PR 21-AUG-1996; 96US-0024279.
XX
PR 02-AUG-1996; 96AU-0001402.
XX
PA (AUST-) AUSTIN RES INST.
XX
PI McKenzie IFC, Sandrin MS;
XX
DR WPI; 1998-159170/14.
XX
DR N-PSDB; AAV21639.
XX
PT Nucleic acids encoding chimeric glycosyltransferases - used for
altering carbohydrate levels on the surface of cells, useful in gene
therapy and transplantation
XX
PS Example 2; Fig 6; 51pp; English.
XX
CC The present sequence represents pig secretor used in an example of the
present invention. The present invention describes nucleic acids (NA)
encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
domain of a first glycosyltransferase (GT) and a localisation signal of
a second GT, whereby when the NA is expressed in a cell and where the
chimeric enzyme is located in an area of the cell where it is able to
compete for substrate with a second GT, resulting in reduced levels of
a product from the second GT. The NAs can be used to produce cells and
organs with desired glycosylation patterns. Products and methods of the
present invention can be used to reduce the levels of undesirable
epitopes in cells, tissues or organs which may be used in
transplantation or gene therapy.
XX
SQ Sequence 340 AA;
Query Match 73.9%; Score 1502; DB 19; Length 340;
Best Local Similarity 79.6%; Pred. No. 9.1e-155;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;
QY 1 MASAOVPFFFPPLAHFLIFVFTSTIIHLQORIVKLOPLSEKELPMTQMSSGNTSPESMR 60
DB 1 MLSMQASFFFPPTGPFLFTASTIIFVFTSTIIHLQORIVKLOPLSEKELPMTQMSSGNTSPESMR 52
QY 61 RSEQHGNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAFIPRI 120
DB 53 -----SSPOLKGMWTINAGRLGNOMGEYATLALARMNGRPAFIPPMHSTLAFIPRI 106
QY 121 SLFVLHSDPAKKIPQWYHLNDWMEERYHIFGCHVRFRTGYPSCSWTFYHHLRPEILKEFT 180
DB 107 TLPVLHASTARRIPQWYHLNDWMEERYHIFGEXYVRLFGYPCSWTFYHHLRTEILREFT 166
QY 181 LHDHVREAAQAFRLGRVNGSQPSTFVGHVHRGDYVHYMPNVNKGVVADRYLEKALDM 240
DB 167 LHNHVREAAQDFRLGRVNGSRPSTYVGHVHRGDYVHYMPNVNKGVVADRYLEQALDW 226
QY 241 FRARYSSPVFVVTSGNMCNRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 300
DB 227 FRARYSPFVTVSSNGMCNRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTIG 286
QY 301 TFGIWAAYLAGGDTIYLANVTLTLPDSPFLKVFKEPEAAFLPEWVGIPADLSPLLK 353
DB 287 TFGIWAAYLAGGETIYLANVTLTLPDSPFLKLFKEPEAAFLPEWIGIEADLSPLLK 339
RESULT 7
AAW69332
ID AAW69332 standard; Protein; 343 AA.
XX
AC AAW69332;
XX
DT 20-NOV-1998 (first entry)
XX
DE Human Sec2 protein sequence.
XX
KW Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
GDP-L-fucose;beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
FUT2; nonsecretor genotyping.
XX
OS Homo sapiens.
XX
PN US5807732-A.
XX
PD 15-SEP-1998.
XX
PF 28-FEB-1995; 95US-0395800.
XX
PR 28-FEB-1995; 95US-0395800.
XX
PA (GIOR/) GIORGI D.
PA (KELL/) KELLY R J.
PA (LENN/) LENNON G.
PA (LOWE/) LOWE J B.
PA (ROUQ/) ROUQUET S.
XX
PI Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
XX
WPI; 1998-520127/44.
XX
DR N-PSDB; AAV58323.
XX
PT DNA encoding fucosyltransferase enzyme - useful for producing
recombinant enzyme and genotyping person as secretor or nonsecretor
XX
PS Disclosure; Column 45-50; 55pp; English.
XX
CC This sequence is the human Sec2 protein of the invention. The DNA
encodes a alpha(1,2) fucosyltransferase and is the Secretor
CC alpha(1,2)fucosyltransferase locus, that cross hybridises with the
CC H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC individual as a secretor or nonsecretor as it is known that nonsecretors
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC the position corresponding to amino acid 143.
XX
SQ Sequence 343 AA;
Query Match 73.1%; Score 1485.5; DB 19; Length 343;
Best Local Similarity 77.9%; Pred. No. 5.8e-153;
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;
QY 1 MASAOVPFFFPPLAHFLIFVFTSTIIHLQORIVKLOPLSEKELPMTQMSSGNTSPESMR 60
DB 1 MLVQMPFSEPAHFLIFVFTSTIIFVQORLAKIQAM--WELPV-----QIPVLA 49
QY 61 RSEQHGNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAFIPRI 120
DB 50 SYSKALGPSQLRGMTWITNAGRLGNOMGEYATLALAKMNGRPAFIPQAMHSTLAFIPRI 109

```

QY 121 SLFVLSHSDAKKIPQWNYHLNDWMEERYRHIGHFVRFTGYPCSWTFYVHHLRPEILKEPT 180
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 110 TLPLVLSHATASRIPQWNYHLNDWMEERYRHIGFVYRFTGYPCSWTFYVHHLRQELQEST 169
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 181 LHDHVREAAQAFRLGLRVNGSQPSTFVGVHVRGDDYVHVMPNVKGVWVADRGLYLEKALDM 240
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 170 LHDHVREAAQAFRLGLRVNGSQPSTFVGVHVRGDDYVHVMPNVKGVWVADRGLYLEKALDM 229
 QY 241 FRARYSSPVFVVTNSGMAWCRENINASRGDVFAGNGIEGSPAKDFALLTOCNHTIMTI 300
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 230 FRARYSSPVFVVTNSGMAWCRENINASRGDVFAGNGIEGSPAKDFALLTOCNHTIMTI 289
 QY 301 TFGIWAAYLAGGDTIYLYANTLPDPSFLKVFKEPAAFLEPWGIPADLSPLLK 353
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 290 TFGIWAAYLTGGDTIYLYANTLPDPSFLKVFKEPAAFLEPWGIPADLSPLLK 342
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 RESULT 8
 AAB36106
 ID AAB36106 standard; Protein: 344 AA.
 AC AAB36106;
 DT 19-FEB-2001 (first entry)
 DE Human Sec2 catalytic domain.
 KW Human; Sec2: alphas-2-fucosyltransferase; cytosolic;
 KW neuroprotective; neurotropic; gene therapy; Fucal-2-galactoside-3-galNAc;
 KW immunotherapy; immunosuppression; cancer; neurological disease;
 KW small cell lung carcinoma.
 XX Homo sapiens.
 XX WC200064464-A1.
 XX 02-NOV-2000.
 XX 23-APR-1999; 99WO-US07384.
 XX 23-APR-1999; 99WO-US07384.
 XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.
 XX Holmes EH, Sherwood AL;
 XX WPI; 2000-687262/67.
 XX New rat ganglioside GM1-specific alphas-2-fucosyltransferase, useful for
 XX preparation of fucosyl GM1 which is useful as a nutritional composition
 XX or immunotherapeutic for cancer and neurological diseases -
 XX Example; Fig 3B; 9lpp; English.
 XX The present sequence is given in a specification relating to a rat
 XX ganglioside GM1-specific alphas-2-fucosyltransferase protein. The protein
 XX or its cellular fraction is useful for synthesis of a molecule comprising
 XX Fucal-2-galactoside-3-galNAc, a glycolipid, glycoprotein, glycolipoprotein
 XX or a free oligosaccharide comprising Fucal-2-galactoside-3-galNAc.
 XX The method involves contacting alphas-2-fucosyltransferase with GDP-fucose
 XX and a molecule or glycolipid, glycoprotein, glycolipoprotein or
 XX oligosaccharide having a terminal galactoside-3-galNAc group. It is also
 XX useful for synthesis of fucosyl-GM1 by contacting the protein with
 XX GDP-fucose and ganglioside GM1. The obtained glycoproteins,
 XX glycolipoproteins, glycolipids and oligosaccharides are useful as
 XX nutritional compositions and fucosyl-GM1 is useful for inducing an
 XX immunotherapeutic or immunosuppressive action against cancer,
 XX neurological disease or small cell lung carcinoma.
 SQ Sequence 344 AA;
 Query Match 72.6%; Score 1475; DB 21; Length 344;
 XX

Best Local Similarity 77.7%; Pred. No. 8.1e-152;
 Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;
 QY 1 MASQVPFSPFLAHFLIFVFTVSTIIHQORIVKLOPLSEKELPMTQMSNGTESPEMR 60
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 1 MLVQVMPFSPFLAHFLIFVFTVSTIIHQORIVKLOPLSEKELPMTQMSNGTESPEMR 49
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 61 RSEQHGNGELGCMPTINSIGRLGNOMGEYATILFALARMNGRLAFIPASMHNALADIPRI 120
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 50 STSKALGPSQLRGWNTINAIGRLGNOMGEYATILFALARMNGRLAFIPASMHNALADIPRI 109
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 121 SLFVLSHSDAKKIPQWNYHLNDWMEERYRHIGHFVRFTGYPCSWTFYVHHLRPEILKEPT 179
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 110 TLPLVLSHATASRIPQWNYHLNDWMEERYRHIGFVYRFTGYPCSWTFYVHHLRQELQEST 169
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 180 TLHDHVREAAQAFRLGLRVNGSQPSTFVGVHVRGDDYVHVMPNVKGVWVADRGLYLEKALD 239
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 170 TLHDHVREAAQAFRLGLRVNGSQPSTFVGVHVRGDDYVHVMPNVKGVWVADRGLYLEKALD 229
 QY 240 MFRARYSSPVFVVTNSGMAWCRENINASRGDVFAGNGIEGSPAKDFALLTOCNHTIMTI 299
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 230 MFRARYSSPVFVVTNSGMAWCRENINASRGDVFAGNGIEGSPAKDFALLTOCNHTIMTI 289
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 300 TFGIWAAYLAGGDTIYLYANTLPDPSFLKVFKEPAAFLEPWGIPADLSPLLK 353
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 QY 290 TFGIWAAYLTGGDTIYLYANTLPDPSFLKVFKEPAAFLEPWGIPADLSPLLK 343
 Db :||||| || :||||| ||||||| ||||||| :||||| ||||||| ||||||| |||||||
 RESULT 9
 AAR13751
 ID AAR13751 standard; Protein: 365 AA.
 XX AAR13751;
 XX 07-NOV-1991 (first entry)
 DE GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase.
 KW Glycosyltransferase.
 OS Homo sapiens.
 XX WO9112340-A.
 XX 22-AUG-1991.
 XX 14-FEB-1991; 91WO-US00899.
 XX 12-DEC-1990; 90US-0627621.
 XX 14-FEB-1990; 90US-0479858.
 XX 14-FEB-1990; 90US-0480133.
 XX (UNMI) UNIV OF MICHIGAN.
 XX Lowe JB;
 XX WPI; 1991-267151/36.
 XX P-PSDB; AAR13751.
 XX Isolation of gene conveying post-translational characteristic -
 XX e.g. the presence of soluble or membrane bound oligo or
 XX polysaccharide or glycosyltransferase.
 XX Disclosure; Fig 3; 155pp; English.
 XX The amino acid sequence codes for a protein capable of functioning
 XX as a GDP-Fuc:beta-D-gal alpha(1,2)-fucosyltransferase. The
 XX functional protein is represented by amino acids 33 to 365. The
 XX enzyme produced by the DNA sequence can be used in enzymatic
 XX fucosylation of chain-terminating galactose residues on lactose-
 XX amine or neolactose type beta-D-galactoside to alpha-2-L-fucose
 XX residues. See also AAR13749-R13752.
 XX

OS Homo sapiens.

XX W09534202-A1.

XX PD 21-DEC-1995.

XX 14-JUN-1995; 95WO-US07554.

XX PR 21-JUL-1994; 94US-0278282.

XX PR 15-JUN-1994; 94US-0260201.

XX PA (ALEX-) ALEXION PHARM INC.

XX PA (AUST-) AUSTIN RES INST.

XX PI Fodor WL, Mckenzie IFC, Rother RP, Sandrin MS, Squinto SP;

XX DR WPI; 1996-049326/05.

XX DR N-PSDB; AAT12238.

XX Redn. of rejection of xenogeneic cells following transplantation
PT by introducing a vector expressing fucosyltransferase into the
PT cells

XX Example 1; Page 45-47; 69pp; English.

XX CC The human H-transferase (AAR90572) product of a cDNA clone (AAT12238)
CC prep'd. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
CC can be expressed in xenogeneic organs, tissues and cells using
CC a vector such as pAPEX-1 (AAT12239). This results in decreased
CC expression of the non-human antigen galactose alpha(1,3) galactose
CC on the surface of the organs etc. so that hyperacute rejection is
CC reduced upon transplantation to humans.

XX SQ Sequence 365 AA;

Query Match 52.0%; Score 1057.5; DB 17; Length 365;

Best Local Similarity 56.5%; Pred. No. 3.1e-106;

Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHQ-----RIVKLQPLSEKELPMTQMSSGNTE 56

DB 11 LAFLVCVLSVIFFLHIHQDSFPHGLGLSILCPDRLVTPPVAFICLPGTAGMPNASSC 70

QY 57 PEMRDSQHONGELRGMTINSTIGRLGNQMGEXATLFPALARMNGRLAFIPASMHNALAP 116

DB 71 P-----QH-PASLSGTWTVYPNGRFGNQMGQVATLLAQLAGRRRAFIIPAMHAALAP 122

QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYRHPGHEVFTGVPSCSWTYHHLRPEIL 176

DB 123 VFRITLPVLAPEVDSRTPWRQLQHDWMESEYADLRDPFLKSGFPCCSWTFPHHLREQIR 182

QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGHVHVRGDIYVHMPNWKGVVADRGYL 234

DB 183 REFTLHDHLREEAQSVLQQLRGTRGDRPRTFVGHVHVRGDIYLVMPQWKGVSAYL 242

QY 235 EKALDMFRARYSSPVVVTNSGMWACRENINASRGDVVFAGNIEGSPAKDFALLTCQNH 294

DB 243 RQAMDWFRARHAPVFTVNSGMWCKENIDTSQDVTFFAGDQGEATPWKDFALLTCQNH 302

QY 295 TIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKVPKEAAFLPEWVGIPADLSPLKA 354

DB 303 TIMTIGTGGWAAYLAGDGTIYLANFTLPDSFELKVPKEAAFLPEWVGIPADLSPLKA 362

QY 355 LTP 357

DB 363 AKP 365

RESULT 14

AAW23805

ID AAW23805 standard; Protein; 365 AA.

XX

AC AAW23805;

XX 15-SEP-1997 (first entry)

XX Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).

XX KW Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
KW alpha 1,2 FT; transgene; transgenic mouse; animal model;
KW intestinal adhesion; Helicobacter pylori infection; stomach;
KW small intestine; gut; epithelial cell; surface receptor;
KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW gastric adenocarcinoma; Lewis antigen; fucosylation.

XX OS Homo sapiens.

XX PN US5625124-A.

XX PD 29-APR-1997.

XX PF 11-JUL-1994; 94US-0273411.

XX PR 11-JUL-1994; 94US-0273411.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Falk P, Gordon JI;

XX DR WPI; 1997-258275/23.

XX DR N-PSDB; AAT17668.

XX Animal model for Helicobacter pylori infection - comprising
PT transgenic mouse expressing human enzyme promoting intestinal
PT adhesion

XX Example 1; Columns 13-16; 24pp; English.

XX A claimed transgenic mouse expresses, in its intestinal epithelial
CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
CC fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
CC called alpha 1,3/4 FT). The enzyme is expressed under the
CC control of a gut epithelial cell-specific promoter and Helicobacter
CC pylori adheres to the transgenic cells. The transgenic mouse and
CC compounds for the ability to inhibit adhesion of H. pylori to gut
CC intestinal epithelial cells from it are useful as models for screening
CC epithelial cells. The present sequence represents human alpha 1,2 FT
CC and was published by Larsen et al., Proc.Natl.Acad.Sci.USA, 87:
CC 6674-6678 (1990).

XX SQ Sequence 365 AA;

Query Match 52.0%; Score 1057.5; DB 18; Length 365;

Best Local Similarity 56.5%; Pred. No. 3.1e-106;

Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHQ-----RIVKLQPLSEKELPMTQMSSGNTE 56

DB 11 LAFLVCVLSVIFFLHIHQDSFPHGLGLSILCPDRLVTPPVAFICLPGTAGMPNASSC 70

QY 57 PEMRDSQHONGELRGMTINSTIGRLGNQMGEXATLFPALARMNGRLAFIPASMHNALAP 116

DB 71 P-----QH-PASLSGTWTVYPNGRFGNQMGQVATLLAQLAGRRRAFIIPAMHAALAP 122

QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYRHPGHEVFTGVPSCSWTYHHLRPEIL 176

DB 123 VFRITLPVLAPEVDSRTPWRQLQHDWMESEYADLRDPFLKSGFPCCSWTFPHHLREQIR 182

QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGHVHVRGDIYVHMPNWKGVVADRGYL 234

DB 183 REFTLHDHLREEAQSVLQQLRGTRGDRPRTFVGHVHVRGDIYLVMPQWKGVSAYL 242

QY 235 EKALDMFRARYSSPVVVTNSGMWACRENINASRGDVVFAGNIEGSPAKDFALLTCQNH 294

DB 243 RQAMDWFRARHAPVFTVNSGMWCKENIDTSQDVTFFAGDQGEATPWKDFALLTCQNH 302

```

QY 295 TIMTIGTGGWAAYLAGGDTIYLANFTLPDSPLKVKFKEAAFLPEWVGIPADLSPLLKA 354
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TIMTIGTGGWAAYLAGGDTIYLANFTLPDSPLKVKFKEAAFLPEWVGINADLSPLWTL 362
QY 355 LTP 357
      |
Db 363 AKP 365

RESULT 15
AAW13640
ID AAW13640 standard; Protein; 365 AA.
AC
XX
XX AAW13640;
XX
DT 19-JUN-1997 (first entry)
XX
DE Human alpha(1,2)-fucosyltransferase.
XX
KW Alpha(1,2)-fucosyltransferase; glycosylation; oligosaccharide;
KW blood group H.
XX
OS Homo sapiens.
XX
PN WC9709421-A1.
XX
PD 13-MAR-1997.
XX
PF 06-SEP-1996; 96WO-US13816.
XX
PR 08-SEP-1995; 95US-0525058.
XX
PA (UNMI ) UNIV MICHIGAN.
XX
PI Legault DJ, Lowe JB;
XX
DR WPI; 1997-192897/17.
DR N-PSDB; AAT61677.
XX
PT New recombinant fucosyltransferase proteins - useful for modifying
PT cell surface oligosaccharide structures
XX
PS Example 1; Page 280-281; 329pp; English.
XX
CC Human GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase
CC (AAW13640) is an enzyme involved in the expression of type I and II
CC blood group H structures. Its amino acid sequence was deduced
CC from an isolated DNA (AAT61677) derived from human A431 cells.
CC Expression of the fucosyltransferase animal cell lines provides
CC specific capabilities with respect to post-translational
CC modification of the oligosaccharides of expressed proteins or
CC lipids. Specific applns. of the enzyme include enzymatic
CC fucosylation of chain-terminating galactose residues on
CC lactosamine or neolacto type beta-D-galactoside to alpha-2-
CC fucose residues. The enzyme can also be used to raise antibodies
CC as diagnostic reagents and to screen cpds. for fucosyltransferase
CC inhibitor activity.
XX
SQ Sequence 365 AA;

```

```

Query Match          52.0%; Score 1057.5; DB 13; Length 365;
Best Local Similarity 56.5%; Pred. No. 3.1e-106;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHLOQ-----RIVKLQPLSEKELPMPTOMSSGNTES 56
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 LAFLVCLVSLVIFFIHODSPFHGLGLSILCPDRRLVTPPVAIFCLPCTAMGNASSC 70
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 PEMRDXSEQHNGELGEMFTINSIGRLGNQMGYATLFLAARMNGRLAIPASMHNALAP 116
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 P-----QH-PASLSGTWTVYPNGRFNGMQGYATLLALAQLNGRRFILLPAMHAALAP 122

```

```

QY 117 IFRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFRFTGYPCSWTFYHHLRPEIL 176
      :|||:|||| : : ||| : ||||| | : | : | : | : | : | : | : | : |
Db 123 VFRTITPLVLAPEVDSPTPRRELQLDHGMSEYADLDKDPFLKLSGFPSCSWTFYHHLRQIR 182
      :|||:||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: ||
QY 177 KEPTLHDHVREERAQAFRLGRV--NGSQPSTFVGVHVRGDYVHVMPNWKGVYVADRGYL 234
      :|||:||||: |||||: ||| : | : ||||| |||||: || ||||| ||
Db 183 REFTLHDHLREEAQSIVLGQLRLGRTGDRPRTFVGVHVRGDYLVQVMPQKWKVVGDSAYL 242
      :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
QY 235 EKALDMFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNH 294
      :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
Db 243 RQAWDFRARHEAPVVFVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNH 302
      :|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||
QY 295 TIMTIGTGGWAAYLAGGDTIYLANFTLPDSPLKVKFKEAAFLPEWVGIPADLSPLLKA 354
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TIMTIGTGGWAAYLAGGDTIYLANFTLPDSPLKVKFKEAAFLPEWVGINADLSPLWTL 362
QY 355 LTP 357
      |
Db 363 AKP 365

```

Search completed: May 27, 2003, 15:06:14
Job time : 52.1951 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 15.6237 Seconds
(without alignments)
715.625 Million cell updates/sec

Title: us-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVFPFPLAHFLIFV.....RSHEHLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

- 1: /cgn2.6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2.6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2.6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2.6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2.6/prodata/1/iaa/PCUS_COMB.pep.*
- 6: /cgn2.6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	380	4	US-09-298-886-8
2	1896	93.3	353	4	US-09-298-886-10
3	1502	73.9	340	4	US-09-254-077A-6
4	1502	73.9	340	4	US-09-254-077A-7
5	1485.5	73.1	343	1	US-08-395-800A-8
6	1485.5	73.1	343	4	US-09-254-077A-8
7	1475	72.6	344	4	US-09-298-886-11
8	1472.5	72.5	347	4	US-09-254-077A-9
9	1086.5	53.5	373	4	US-09-254-077A-12
10	1057.5	52.0	365	1	US-07-914-281-6
11	1057.5	52.0	365	1	US-08-393-246-6
12	1057.5	52.0	365	1	US-08-273-411-1
13	1057.5	52.0	365	1	US-08-525-058A-6
14	1057.5	52.0	365	1	US-08-395-800A-6
15	1057.5	52.0	365	1	US-08-395-800A-10
16	1057.5	52.0	365	2	US-08-696-731-6
17	1057.5	52.0	365	4	US-09-042-531-6
18	1057.5	52.0	365	4	US-09-254-077A-11
19	1057.5	52.0	365	5	PCN-US91-00899-12
20	1051.5	51.7	333	5	PCN-US91-00899-11
21	1046.5	51.5	365	4	US-09-151-592-2
22	1046.5	51.5	365	4	US-09-254-077A-10
23	722.5	35.6	222	1	US-08-395-800A-2
24	439	21.6	102	1	US-08-395-800A-3
25	141	6.9	300	4	US-09-433-598-2
26	97	4.8	183	1	US-08-167-035-33
27	97	4.8	183	1	US-08-208-887A-33

ALIGNMENTS

RESULT 1

US-09-298-886-8

; Sequence 8, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE OF INVENTION: THEREOF

; CURRENT APPLICATION NUMBER: US/09/298,886

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 8

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-298-886-8

Query Match 100.0%; Score 2032; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 3.1e-226;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAQVFPFPLAHFLIFVFTSTIIHQRIYKVLQPLSEKELPMTTOMSGNTESPEMR	50
Db	1	MASAQVFPFPLAHFLIFVFTSTIIHQRIYKVLQPLSEKELPMTTOMSGNTESPEMR	50
QY	61	RDSHGHNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASHMNALAPIRI	120
Db	61	RDSHGHNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASHMNALAPIRI	120
QY	121	SLPVLHSDTAKKIPQNYHLNDWMEERYRHPGHFVFTGYPCSWTFYHHLRPEILKEFT	180
Db	121	SLPVLHSDTAKKIPQNYHLNDWMEERYRHPGHFVFTGYPCSWTFYHHLRPEILKEFT	180
QY	181	LHDVREEAQAFRLGLRVNGSQSTFVGVHVRGDYVHVMPNVKGVVADRGYLEKALDM	240
Db	181	LHDVREEAQAFRLGLRVNGSQSTFVGVHVRGDYVHVMPNVKGVVADRGYLEKALDM	240
QY	241	FRARYSPVFWVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTOCHNTIMTIG	300
Db	241	FRARYSPVFWVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTOCHNTIMTIG	300
QY	301	TFGIWAAYLAGGDTIYLANVTLFDSPLKVKFEAAFLPEWVGIFADLSPLLKALFPACP	360
Db	301	TFGIWAAYLAGGDTIYLANVTLFDSPLKVKFEAAFLPEWVGIFADLSPLLKALFPACP	360
QY	361	RSHEHLKAKGVTCYVAGRAF	380

```
D6 361 RSHFELKAKGVTCYVAGRAF 380

RESULT 2
US-09-298-886-10
; Sequence 10, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-298-886-10

Query Match 93.3%; Score 1896; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.5e-210;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQORIVKLOPSEKELPMTQSSGNTSPENRRDSEQHNGELRGMTINSIGRLGNOM 87
DB 1 LQORIVKLOPSEKELPMTQSSGNTSPENRRDSEQHNGELRGMTINSIGRLGNOM 60

QY 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSTAKKIPQNYHLNDWMEER 147
DB 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSTAKKIPQNYHLNDWMEER 120

QY 148 YRHIPGHVRFYTGPCSWTFYHHLRPEILKEPTLHDHVREAAQALRGLRVNGSQPSTFV 207
DB 121 YRHIPGHVRFYTGPCSWTFYHHLRPEILKEPTLHDHVREAAQALRGLRVNGSQPSTFV 180

QY 208 GVHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCENINAS 267
DB 181 GVHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCENINAS 240

QY 268 RGDVVFAGNIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGSDTIYLANITLPDSPF 327
DB 241 RGDVVFAGNIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGSDTIYLANITLPDSPF 300

QY 328 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 380
DB 301 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 353

RESULT 3
US-09-254-077A-6
; Sequence 6, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.SUSWO
; CURRENT FILING DATE: 1999-06-11
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-6

Query Match 73.9%; Score 1502; DB 4; Length 340;
Best Local Similarity 79.6%; Pred. No. 5.1e-165;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPSFFPLAHLFVFTVSTIIHLQORIVKLOPSEKELPMTQSSGNTSPENMR 60
DB 1 MLQMQAASFFPGPILFVFTASTIFHLQORIVKLOP--TWELQMVQTV---TESP--- 52

QY 61 RDSEHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 120
DB 53 -----SSPQLKGMWTINAIGRLGNOMGEYATLALARMNGRPAFIPPEMHTLAPIRI 106

QY 121 SLPLVLSHTAKKIPQNYHLNDWMEERYRHIPGHVRFYTGPCSWTFYHHLRPEILKEPT 180
DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGEYVRLTYGTPCSWTFYHHLRTEILREFT 166

QY 181 LHDHVREAAQALRGLRVNGSQPSTFVGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
DB 167 LHNHVREAAQDLRLGLRVNGSRPSTYGVHVRRGDYVHVMPNVKGVVADRRYLEQALDW 226

Query Match 73.9%; Score 1502; DB 4; Length 340;
Best Local Similarity 79.6%; Pred. No. 5.1e-165;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAQVPSFFPLAHLFVFTVSTIIHLQORIVKLOPSEKELPMTQSSGNTSPENMR 60
DB 1 MLQMQAASFFPGPILFVFTASTIFHLQORIVKLOP--TWELQMVQTV---TESP--- 52

QY 61 RDSEHGNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFR 120
DB 53 -----SSPQLKGMWTINAIGRLGNOMGEYATLALARMNGRPAFIPPEMHTLAPIRI 106

QY 121 SLPLVLSHTAKKIPQNYHLNDWMEERYRHIPGHVRFYTGPCSWTFYHHLRPEILKEPT 180
DB 107 TLPVLHASTARRIPQNYHLNDWMEERYRHIPGEYVRLTYGTPCSWTFYHHLRTEILREFT 166

QY 181 LHDHVREAAQALRGLRVNGSQPSTFVGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
DB 167 LHNHVREAAQDLRLGLRVNGSRPSTYGVHVRRGDYVHVMPNVKGVVADRRYLEQALDW 226
```

```

QY 241 FRARYSPFVVVTNGMAWORENTNARGDGVVVFAGNIESSPAKDFALLTCQNTHTMTIG 300
Db 227 FRARYSPFVVYSSNGMAWCORENTNARGDGVVVFAGNIESSPAKDFALLTCQNTHTMTIG 286
QY 301 TFGWAAVLAGGDIYLIANYTLDPSPFLKVPKPAFLPEWVGIPADLSPLK 353
Db 287 TFGWAAVLAGGEIYLIANYTLDPSPFLKVPKPAFLPEWIGIEADLSPLK 339

RESULT 5
US-08-395-800A-8
: Sequence 8, Application US/08395800A
: Patent No. 5807732
: GENERAL INFORMATION:
: APPLICANT: LOWE, JOHN B
: APPLICANT: LENNON, GREGORY
: APPLICANT: ROQUIER, SYLVIE
: APPLICANT: GIORGI, DOMINIQUE
: APPLICANT: KELLY, ROBERT J
: TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
: TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
: TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
: TITLE OF INVENTION: GENOTYPING A PERSON
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/395,800A
: FILING DATE: 28-FEB-1995
: CLASSIFICATION: 435
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 343 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-395-800A-8

Query Match 73.1%; Score 1485.5; DB 1; Length 343;
Best Local Similarity 77.9%; Pred. No. 4.2e-163;
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;

QY 1 MASAQVPFSPFLAHLFIYFVTTIIHLQRIQVVKIQPLSEKELPMTTOMSSGNTESPEMR 60
Db 1 MLVQMPFSPMAHFIYFVTVSTFIHVQRLAKIQAM--WELDV-----QIPVLA 49
QY 61 RDSQHGNGELRGMTINSTIGRLNQMGEXATIFALARMNGRLAIFPASHMNALAPIFRI 120
Db 50 STSKALGPSQLRGMTINAIIGRLNQMGEXATLYALAKMGRPAFIQOMHSTLAPIFRI 109
QY 121 SLPLVHSDTAKKIPQNYHLNDWMEERYRHPGHVFRTGYPCSWTFYHHLRPEILKEFT 180
Db 110 TLPVHLSATASRIPQNYHLNDWMEERYRHPGHVFRTGYPCSWTFYHHLRQILQEFT 169
QY 181 LHDHVRREAQAFRLGRVNGSOPSTFVGWVRRGDYVHVMPVWKGVVADRGYLEKALDM 240
Db 170 LHDHVRREAQKFLRGLQVNGSRPGTFVGWVRRGDYVHVMPVWKGVVADRRYLDQQLDW 229

```



```
Db 243 ROAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQCEATPWKDFALLTQCNIH 302
QY 295 TINTICTGFWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLSPLLLKA 354
Db 303 TINTICTGFWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLSPLLLKA 362
QY 355 LTP 357
Db 363 AKP 365

RESULT 12
US-08-273-411-1
; Sequence 1, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; DISEASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..365
; OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365
US-08-273-411-1

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;
QY 12 LAHFLFFVFTVITHLQ-----RIVKLQPLSEKELPMTTQSSGNIES 56
```

```
Db 11 LAFLVCLSVIFFELIHODSFFPHGLGLSLQCDRLVTPPPVAIFCLPGTAMGNASSSC 70
QY 57 PEMRSESEONGELRGMTINTISIGRLGNQMGYAILFALARMNGRLAIPASMHNALAP 116
Db 71 P-----QH-PASLSGTVTVPNGRFGNQMGQYATLLAQLNGRRATILFAMHAALAP 122
QY 117 IFRISLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHEVFRFTGYPCSWTFYHHLRPEIL 176
Db 123 VFRITLPVLAPEVDSTFPRRELQLDHDMSEYADLRDPFLKLSGFCSTFFHHLREQIR 182
QY 177 KEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGVHVRKGDYVHVMPNWKGVYVADRGYL 234
Db 183 REFTLHDHLKEEAQSVGLQRLRGRTGDRPTFVGVHVRKGDYLVQVMPQKVGKVGVSAYL 242
QY 235 EKALDMFRARYSSPVFVYVTSNGNAWCRENINASRGDVTAGNGIEGSPAKDFALLTQCNIH 294
Db 243 ROAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQCEATPWKDFALLTQCNIH 302
QY 295 TINTICTGFWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLSPLLLKA 354
Db 303 TINTICTGFWAAYLAGGDTIYLANVTLDPSPFLKVKPEAAFLPEWVGIPADLSPLLLKA 362
QY 355 LTP 357
Db 363 AKP 365

RESULT 13
US-08-525-058A-6
; Sequence 6, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LORE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-4500
; TELEFAX: (703) 486-2347
; TELETYPE: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-058A-6

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
```


; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-395-800A-10

Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-113;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY	12	LAHFLIFVFTSTIIHQ-----RIVKLOPLSEKELPMTQMSSGNIES	56
DB	11	LAFLVCLVLSVIFELHIHQDSFPHGLGLSILCPDRRLVTPPVAIFCLPGTAMGNASSC	70
QY	57	PEMRDSEQHONGELRGFTTNSIGRLGNQMGYATLFLALPMNGRLAFIPASMHNALAP	116
DB	71	P-----QH-PASLSGTWTVYPNGRFGNQMGYATLIALAQLNGRRAFILFAMHAALAP	122
QY	117	IFRISLPVLHSDTAKKIPQWQNYHLNDMEERYRHIPGHFVRFYPCSWTFYHHLRPEIL	176
DB	123	VFRITLPVLAPEVDSRTFWRLEQLHDWSEYADLRDPFLKLSGFPSCSWTFHHLREQIR	182
QY	177	KEFTLHDHVRREAQAFRLGLRV--NGSQPSTFVGHVHRRGDYVHVMNVKGVVADRQYL	234
DB	183	REFTLHDHLREEAQSVLQGLRLGRTGDRPRTFVGHVHRRGDYLVQMPQRKGVVGDQAYL	242
QY	235	EKALDMFRARYSSPVVYTSNGMAWCRENINASRGDVVFNAGNIEGSPAKDFALLTQCNH	294
DB	243	RQAMDWFRARHEAPVVFVTSNGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALLTQCNH	302
QY	295	TIMTIGTFGIWAAYLAGDITIYLANITLPDSPLKVKPEAAFLPEWVGIPADLSPLKA	354
DB	303	TIMTIGTFGIWAAYLAGDVIYLANITLPDSEFLKIFKPEAAFLPEWVGINADLSPLWL	362
QY	355	LTP	357
DB	363	AKP	365

Search completed: May 27, 2003, 15:11:50
Job time : 17.6237 secs

GenCore version: 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:08:29 ; Search time 16.4181 Seconds
(without alignments)
2295.367 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAQVPFSFPLAHFLIFV.....RSHEHLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	380	10	US-09-999-672-8
2	2032	100.0	380	12	US-10-040-863-8
3	1896	93.3	353	10	US-09-999-672-10
4	1896	93.3	353	12	US-10-040-863-10
5	1502	73.9	340	10	US-09-051-034A-2
6	1475	72.6	344	10	US-09-999-672-11
7	1475	72.6	344	12	US-10-040-863-11
8	1057.5	52.0	365	9	US-10-105-963-10
9	1057.5	52.0	365	10	US-09-863-475A-6
10	1046.5	51.5	365	10	US-09-051-034A-4
11	1044.5	51.4	365	10	US-09-844-268-13
12	1044.5	51.4	365	10	US-09-844-705-13
13	141	6.9	300	10	US-09-848-838-2
14	113	5.6	431	9	US-09-479-614-14
15	113	5.6	496	9	US-09-614-2
16	110.5	5.4	496	9	US-10-214-524-25
17	109	5.4	496	9	US-09-479-614-29
18	89	4.4	1031	9	US-09-764-868-647
19	89	4.4	2057	9	US-09-815-379-10

20	89	4.4	2058	9	US-09-815-379-17	Sequence 17, Appl
21	87	4.3	217	10	US-09-765-298A-6	Sequence 6, Appl
22	85.5	4.2	738	9	US-10-235-521-1	Sequence 1, Appl
23	84.5	4.2	344	9	US-09-895-913A-56	Sequence 56, Appl
24	84.5	4.2	344	10	US-09-815-242-11366	Sequence 11366, A
25	83.5	4.1	747	9	US-10-029-072-212	Sequence 212, App
26	83.5	4.1	747	9	US-10-121-049-212	Sequence 212, App
27	83.5	4.1	747	9	US-10-123-904-212	Sequence 212, App
28	83.5	4.1	747	9	US-10-140-470-212	Sequence 212, App
29	83.5	4.1	747	9	US-10-175-746-212	Sequence 212, App
30	83.5	4.1	747	9	US-10-176-918-212	Sequence 212, App
31	83.5	4.1	747	9	US-10-176-921-212	Sequence 212, App
32	83.5	4.1	747	9	US-10-137-865-212	Sequence 212, App
33	83.5	4.1	747	9	US-10-140-474-212	Sequence 212, App
34	83.5	4.1	747	9	US-10-142-431-212	Sequence 212, App
35	83.5	4.1	747	9	US-10-143-114-212	Sequence 212, App
36	83.5	4.1	747	9	US-10-140-002-212	Sequence 212, App
37	83.5	4.1	747	9	US-10-142-419-212	Sequence 212, App
38	83.5	4.1	747	9	US-10-123-262-212	Sequence 212, App
39	83.5	4.1	747	9	US-10-142-423-212	Sequence 212, App
40	83.5	4.1	747	9	US-10-121-050-212	Sequence 212, App
41	83.5	4.1	747	9	US-10-141-755-212	Sequence 212, App
42	83.5	4.1	747	9	US-10-143-032-212	Sequence 212, App
43	83.5	4.1	747	9	US-10-123-108-212	Sequence 212, App
44	83.5	4.1	747	9	US-10-123-236-212	Sequence 212, App
45	83.5	4.1	747	9	US-10-123-261-212	Sequence 212, App

ALIGNMENTS

RESULT 1
US-09-999-672-8
; Sequence 8, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-8

Query Match	100.0%	Score	2032	DB	10	Length	380
Best Local Similarity	100.0%	Pred. No.	1.6e-200	Mismatches	0	Indels	0
Matches	380	Conservative	0	0	Gaps	0	0
QY	1	MASAQVPFSFPLAHFLIFVVTSTIIHLQQRIVKLPQPLSEKELPMTQSSNGTSPSPMR	60				
Db	1	MASAQVPFSFPLAHFLIFVVTSTIIHLQQRIVKLPQPLSEKELPMTQSSNGTSPSPMR	60				
QY	61	RSEQHGNGELRGMTFINSIGRLGNOMGEYATLFLARNMNGRLAFIPASMHNALAFIPRI	120				
Db	61	RSEQHGNGELRGMTFINSIGRLGNOMGEYATLFLARNMNGRLAFIPASMHNALAFIPRI	120				
QY	121	SLPVLHSDTAKKIPQNYHLNDWMERYRHFHGFVRFTGYPCSWTFYHHLPEILKFT	180				
Db	121	SLPVLHSDTAKKIPQNYHLNDWMERYRHFHGFVRFTGYPCSWTFYHHLPEILKFT	180				
QY	181	LHDHYREEAAQFLRGLRVNGSQPSTFVGHVRRGDYVHVMPNVNKGVSVDARGYLEKALDM	240				
Db	181	LHDHYREEAAQFLRGLRVNGSQPSTFVGHVRRGDYVHVMPNVNKGVSVDARGYLEKALDM	240				

```

QY 241 FRARYSSPVVVTSGMAWCENINASRGDGVVFNAGNIEGSPAKDFALLTQCNTHTMTIG 300
Db 241 FRARYSSPVVVTSGMAWCENINASRGDGVVFNAGNIEGSPAKDFALLTQCNTHTMTIG 300
QY 301 TFGIWAAYLAGDIIYLANYTLPDSPFLKVKPEAAFLPEWVGIPADLSPLLKALTPACP 360
Db 301 TFGIWAAYLAGDIIYLANYTLPDSPFLKVKPEAAFLPEWVGIPADLSPLLKALTPACP 360
QY 361 RSHFHLKAKGVTCYVAGRAF 380
Db 361 RSHFHLKAKGVTCYVAGRAF 380

RESULT 2
US-10-040-863-8
; Sequence 8, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-8

```

```

Query Match 100.0%; Score 2032; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.6e-200;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAOVPSFPLAHFLIEFVFTSTIIHLQQRIVKLOPLSEKELPMTTQSSGNTSPEMR 60
Db 1 MASAOVPSFPLAHFLIEFVFTSTIIHLQQRIVKLOPLSEKELPMTTQSSGNTSPEMR 60
QY 61 RDSQHGNGELRGMTINSIGRLGNQMGYEATLALARMNGRLAFIPASMHNALAPIFRI 120
Db 61 RDSQHGNGELRGMTINSIGRLGNQMGYEATLALARMNGRLAFIPASMHNALAPIFRI 120
QY 121 SLPLVLSHTAKKIPWNYHLNDWMEERYHHPGHVREFTGYPCSWTFYHHLRPEILKFT 180
Db 121 SLPLVLSHTAKKIPWNYHLNDWMEERYHHPGHVREFTGYPCSWTFYHHLRPEILKFT 180
QY 181 LHDHVREAAQFLRGLRVNGSQPSTFVGHVHVRGDDYHVMPNVKGVVADRGYLEKALDM 240
Db 181 LHDHVREAAQFLRGLRVNGSQPSTFVGHVHVRGDDYHVMPNVKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVVVTSGMAWCENINASRGDGVVFNAGNIEGSPAKDFALLTQCNTHTMTIG 300
Db 241 FRARYSSPVVVTSGMAWCENINASRGDGVVFNAGNIEGSPAKDFALLTQCNTHTMTIG 300
QY 301 TFGIWAAYLAGDIIYLANYTLPDSPFLKVKPEAAFLPEWVGIPADLSPLLKALTPACP 360
Db 301 TFGIWAAYLAGDIIYLANYTLPDSPFLKVKPEAAFLPEWVGIPADLSPLLKALTPACP 360
QY 361 RSHFHLKAKGVTCYVAGRAF 380
Db 361 RSHFHLKAKGVTCYVAGRAF 380

```

```

RESULT 3
US-09-999-672-10
; Sequence 10, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-10

Query Match 93.3%; Score 1896; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-186;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQQRIVKLOPLSEKELPMTTQSSGNTSPEMRRDSEQHGNGELRGMTINSIGRLGNOM 87
Db 28 LQQRIVKLOPLSEKELPMTTQSSGNTSPEMRRDSEQHGNGELRGMTINSIGRLGNOM 87
QY 88 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHTAKKIPWNYHLNDWMEER 147
Db 88 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHTAKKIPWNYHLNDWMEER 147
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHTAKKIPWNYHLNDWMEER 120
Db 61 GEYATLALARMNGRLAFIPASMHNALAPIFRIISLPVLSHTAKKIPWNYHLNDWMEER 120
QY 148 YRHIPGHVREFTGYPCSWTFYHHLRPEILKFTLHDHVREAAQFLRGLRVNGSQPSTFV 207
Db 148 YRHIPGHVREFTGYPCSWTFYHHLRPEILKFTLHDHVREAAQFLRGLRVNGSQPSTFV 207
QY 208 GYHVRGDDYHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCENINAS 267
Db 208 GYHVRGDDYHVMPNVKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCENINAS 267
QY 268 RGVVFNAGNIEGSPAKDFALLTQCNTHTMTIGTGIWAAYLAGDIIYLANYTLPDSPF 327
Db 268 RGVVFNAGNIEGSPAKDFALLTQCNTHTMTIGTGIWAAYLAGDIIYLANYTLPDSPF 327
QY 328 LKVFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 380
Db 328 LKVFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 380

RESULT 4
US-10-040-863-10
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

Query Match 93.3%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-186;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQQRIVKLOPLSEKELPMTTQSSGNTSPEMRRDSEQHGNGELRGMTINSIGRLGNOM 87

```

Db 1 LQORIVKLOPLSEKELPMITQSSNGESESERDSEHQHNGELRCGMFTINSIGRLGNQM 60
QY 88 GEYATLFLALARMNGRLAFIPASMHNLALAFIRISLVLSDTAKKIPWQNYHLNDWMEER 147
Db 61 GEYATLFLALARMNGRLAFIPASMHNLALAFIRISLVLSDTAKKIPWQNYHLNDWMEER 120
QY 148 YRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSPSTFV 207
Db 121 YRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSPSTFV 180
QY 208 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 267
Db 181 GVHVRGDIYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINAS 240
QY 268 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLANITLDPDSPF 327
Db 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLANITLDPDSPF 300
QY 328 LKVFKEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 380
Db 301 LKVFKEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 5

US-09-051-034A-2

; Sequence 2, Application US/09051034A

; Patent No. US2001005584A1

; GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL

; APPLICANT: SANDRIN, MAURO SERGIO

; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

; TITLE OF INVENTION: GLYCOSYLTRANSFERASE

; FILE REFERENCE: 30562.6USWO

; CURRENT APPLICATION NUMBER: US/09/051,034A

; CURRENT FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: PCT/AU97/00492

; PRIOR FILING DATE: 1997-08-01

; PRIOR APPLICATION NUMBER: 60/024,279

; PRIOR FILING DATE: 1996-08-21

; PRIOR APPLICATION NUMBER: P01402

; PRIOR FILING DATE: 1996-08-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus Domesticus

US-09-051-034A-2

Query Match 73.9%; Score 1502; DB 10; Length 340;
Best Local Similarity 79.6%; Pred. No. 4.5e-146;
Matches 281; Conservative 28; Mismatches 30; Indels 14; Gaps 3;

QY 1 MASAOVPFSPFLAHFLIFVFTSTIIHLQORIVKLOPLSEKELPMTQSSNGTESPEMR 60
Db 1 MLSMOASFFFTGPFLFVETASTIFHLQORVVKIQP--TWELQVVTQVT---TESP--- 52
QY 61 RDEQHGNGELGCMFTINSIGRLGNOMGEYATLALARMNGRLAFIPASMHNLALAFIRI 120
Db 53 -----SSPOLKGMWMTINAIGRLGNOMGEYATLYALARMNGRPAFIPPEMHSTLAFIRI 106
QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEFT 180
Db 107 TLPLVHASTARIPWQNYHLNDWMEERYRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEFT 166
QY 181 LHDHVREAAQAFRLGLRVNGSQSPSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALDM 240
Db 167 LHNHVREAAQDFLGLRVNGSRPSTVGVHVRGDIYVHVPNVKGVVADRYLEQALDM 226
QY 241 FRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 300
Db 227 FRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTIG 286

QY 301 TFGIWAAYLAGGDTIYLANITLDPDSPFLKVFKEAAFLPEWVGIPADLSPLLK 353
Db 287 TFGIWAAYLAGGDTIYLANITLDPDSPFLKVFKEAAFLPEWVGIPADLSPLLK 339
RESULT 6
US-09-999-672-11
; Sequence 11, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: ERIC H. HOLMES ET AL.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-999-672-11

Query Match 72.6%; Score 1475; DB 10; Length 344;
Best Local Similarity 77.7%; Pred. No. 2.7e-143;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MASAOVPFSPFLAHFLIFVFTSTIIHLQORIVKLOPLSEKELPMTQSSNGTESPEMR 60
Db 1 MLVYQMPFSPFAHFLIFVFTSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
QY 61 RDEQHGNGELGCMFTINSTGRIGNOMGEYATLALARMNGRLAFIPASMHNLALAFIRI 120
Db 50 STSKALGPSOLRCGMWMTINAIGRLGNOMGEYATLYALAKMNGRPAFIPQMHSTLAFIRI 109
QY 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEF 179
Db 110 TLPLVHASTARIPWQNYHLNDWMEERYRHIPGHVFRFTGYPGCSWTFYHHLRPEILKEF 169
QY 180 TLHDHVREAAQAFRLGLRVNGSQSPSTFVGVHVRGDIYVHVPNVKGVVADRGYLEKALD 239
Db 170 TLHDHVREAAQAFRLGLRVNGSRPSTFVGVHVRGDIYVHVPNVKGVVADRYLQOALD 229
QY 240 MFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTHTMTI 299
Db 230 WFRARYSSLIFVTSNGMAWCRENIDTSHGDVVFAGNGIEGSPAKDFALLTQCNTHTMTI 289
QY 300 GTFGIWAAYLAGGDTIYLANITLDPDSPFLKVFKEAAFLPEWVGIPADLSPLLK 353
Db 290 GTFGIWAAYLAGGDTIYLANITLDPDSPFLKVFKEAAFLPEWVGIPADLSPLLK 343

RESULT 7

US-10-040-863-11

; Sequence 11, Application US/10040863

; Patent No. US20020137165A1

; GENERAL INFORMATION:

; APPLICANT: ERIC H. HOLMES ET AL.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/10/040,863

; CURRENT FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: 09/298,886

; PRIOR FILING DATE: 1999-04-23

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0


```
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-863-11

Query Match          72.6%; Score 1475; DB 12; Length 344;
Best Local Similarity 77.7%; Pred. No. 2.7e-143;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MASAOVPSFPLAHFLIEVFVSTIIHQQRIKQIVKLOPLSEKELPMTQMSSGNTESPWR 60
DB 1 MLVQMPSPFMAHFLIEVFVSTIIHQQRIKQIVKLOPLSEKELPMTQMSSGNTESPWR 60
QY 61 KDSQHGNGELRGMTINSIGRLGNQMEYATLFLALAMNGRLAFIPASMHNALAPIRI 120
DB 50 STSKALGPSQLRGMTINAIGRLGNQMEYATLALAKMNGRPAFIPQMSTLAPIRI 109
QY 121 SLPLVHSDTAKIPQNYHLNDWMEERYHTI-PGHVFTGYPCSWTFYHHLRPIKRF 179
DB 110 TLPVLHSDTASRIPQNYHLNDWMEERYHTI-PGHVFTGYPCSWTFYHHLRPIKRF 169
QY 180 TLHDHVRREAQAFRLGLRVNGSPTFVGVHVRGDYVHVMPVWKGVVADRGYLEKALD 239
DB 170 TLHDHVRREAQAFRLGLRVNGSPTFVGVHVRGDYVHVMPVWKGVVADRGYLEKALD 229
QY 240 MFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTI 299
DB 230 MFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNTMTI 289
QY 300 GTFGWAAYLAGGDTIYLANVTLDPSPFLKVPKPEAFLEPWGIPADLSPLLK 353
DB 290 GTFGWAAYLAGGDTIYLANVTLDPSPFLKVPKPEAFLEPWGIPADLSPLLK 343

RESULT 8
US-10-105-963-10
; Sequence 10, Application US/10105963
; Publication No. US20030068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; APPLICANT: Schiff, J. Michael
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; PRIOR FILING DATE: 2002-03-21
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-10

Query Match          52.0%; Score 1057.5; DB 9; Length 365;
Best Local Similarity 56.5%; Pred. No. 2.7e-100;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHQQ-----RIVKLQPLSEKELPMTQMSSGNTES 56
DB 11 LAFLVCVLSVIFVHIIHQDSFPHGLSILCPDRILVTPVVAIFCLGTAMGPASSSC 70
QY 57 PMRRDSOHGNGELRGMTINSIGRLGNQMEYATLFLALAMNGRLAFIPASMHNALAP 116
DB 71 P-----QH-PASLSGTVVTVNCRFGNQGYATLLAQLNGFRRAFILPAMHAALAP 122
QY 117 IFRISLPVHSDTAKIPQNYHLNDWMEERYHTI-PGHVFTGYPCSWTFYHHLRPIEL 176
```

```
Db 123 VFRITPLVLAPEVDSRTPWRELQHLDMSEYADLRDPFLKLSGPPCSWTFPHLREQIR 182
QY 177 KEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGVHVRGDYVHVMPVWKGVVADRGYL 234
DB 183 REFTHLHDHVRREAQAFRLGLRV--NGSOPSTFVGVHVRGDYVHVMPVWKGVVADRGYL 242
QY 235 EKALDMFRARYSSPVFVTSNGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTQCNT 294
DB 243 RQANDWFRARHEAPVFTVTSNGMEWCKENITDSQDVTFAAGDQGAIFKDKFALLTQCNT 302
QY 295 TIMTGTGFWAAYLAGGDTIYLANVTLDPSPFLKVPKPEAFLEPWGIPADLSPLKA 354
DB 303 TIMTGTGFWAAYLAGGDTIYLANVTLDPSPFLKVPKPEAFLEPWGIPADLSPLKA 362
QY 355 LTP 357
DB 363 AKP 365

RESULT 9
US-09-863-475A-6
; Sequence 6, Application US/09863475A
; Patent No. US20020102688A1
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/863,475A
; FILING DATE: 24-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248955 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-863-475A-6

Query Match          52.0%; Score 1057.5; DB 10; Length 365;
Best Local Similarity 56.5%; Pred. No. 2.7e-100;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHQQ-----RIVKLQPLSEKELPMTQMSSGNTES 56
```

```

Db 11 LAFLVGVLSVIFLHIHQDSFPHGLGLSILCPDRRLVTPPVAFICLPGIAMGNASSC 70
Qy 57 PEMRSDSEHQHNGELRGMTFINSIGRLGNQMGVATLAFALARMNGRLAFIPASMHNLAP 116
Db 71 P-----QH-PASLSTWTVYVNGRGNQMGVATLALAQNLGRAFLIPAMHAALAP 122
Qy 117 IFRISLPLVHSDTAKKIPWQNYHLNDWMEERYHIPGHFVFRFTGYPGCSWTFYHHLRPEIL 176
Db 123 VFRITLVLAPVDSRTFWRLEQLHDWMESEYADLRDPFLKLSOFFPCSWTFHHLRQIR 182
Qy 177 KEFTLHDHVRBEAQAFILGRV--NGSQPSTFVGVHVRGGDYVHVMNVKGVVADRGYL 234
Db 183 REFTLHDHVRBEAQSVLQRLGRIGRPRFTFVGVHVRGGDYLVMPQRMKGVVGD SAYL 242
Qy 235 EKALDMPFARYSSVFFVYTSNGMAWCENINASRGDVVFAGNGEGSPAKDFALLTQCNH 294
Db 243 ROADMWFRARBEAFVFFVYTSNGMEWCKENIDTSQDVTTFAGDQGEATPWKDFALLTQCNH 302
Qy 295 TIMIGTGFIAAYLAGDITILANYTLPSDFPLKVKFKEAAFLPEWGPADLSPLIKA 354
Db 303 TIMIGTGFIAAYLAGDITVILANFTLPDSEFLKIKFPEARFLPEWGINADLSPLWTL 362
Qy 355 LTP 357
Db 363 AKP 365

```

RESULT 10

```

US-09-051-034A-4
; Sequence 4, Application US/09051034A
; Patent No. US20010055584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL
; APPLICANT: SANDRIN, MAURO SERGIO
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; FILE REFERENCE: GLYCOSYLTRANSFERASE
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051.034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P01402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Sus Domesticus
US-09-051-034A-4

```

```

Query Match 51.5%; Score 1046.5; DB 10; Length 365;
Best Local Similarity 62.6%; Pred. No. 3.6e-99;
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

```

```

Qy 44 PMTQMSSGNTESPEMRDSEHQHNGELRGMTFINSIGRLGNQMGVATLAFALARMNGRL 103
Db 51 PVAIFCLAGTVPVHPNADSCPKH-PASESGTWITIPDGRFGNQMGVATLALAQNLGRQ 109
Qy 104 AFIPASMHNLAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYHIPGHFVFRFTGYPG 163
Db 110 AFIQPAMHNAVLPVFRITLPVLAPEVDHAPWRELEHDMKSEYAHLEKPEWKLGTGPPC 169
Qy 164 SWTFYHHLRPEILKEFTLHDHVRBEAQAFILGRV--NGSQPSTFVGVHVRGGDYVHVM 221
Db 170 SWTFHHLRQIRSEFTLHDHVRBEAQAFILGRV--NGSQPSTFVGVHVRGGDYVHVM 229
Qy 222 NWKGVVADRGYLEKALDMFRARYSSPVVYTSNGMAWCENINASRGDVVFAGNGIEGS 281
Db 230 KRWKGVVGDGYLQOADMWFRARYEAPVFFVYTSNGMEWCKENIDTSRGDVFAGNGREA 289

```

```

Qy 282 PAKDFALLTQCNTHTIMTIGTGFIAAYLAGDITVILANYTLPSDFPLKVKFKEAAFLPEW 341
Db 290 PARDFALLVQCNTHTIMTIGTGFIAAYLAGDITVILANYTLPSDFPLKVKFKEAAFLPEW 349
Qy 342 VGIADLSPL 351
Db 350 VGINADLSPL 359

```

RESULT 11

```

US-09-844-268-13
; Sequence 13, Application US/09844268
; Patent No. US20020129395A1
; GENERAL INFORMATION:
; APPLICANT: BOSWORTH, BRAD
; APPLICANT: VOGELI, PETER
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
; FILE REFERENCE: 21419/90368
; CURRENT APPLICATION NUMBER: US/09/844,268
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 09/443,766
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Porcine
US-09-844-268-13

```

```

Query Match 51.4%; Score 1044.5; DB 10; Length 365;
Best Local Similarity 62.6%; Pred. No. 5.8e-99;
Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

```

```

Qy 44 PMTQMSSGNTESPEMRDSEHQHNGELRGMTFINSIGRLGNQMGVATLAFALARMNGRL 103
Db 51 PVAIFCLAGTVPVHPNADSCPKH-PASESGTWITIPDGRFGNQMGVATLALAQNLGRQ 109
Qy 104 AFIPASMHNLAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYHIPGHFVFRFTGYPG 163
Db 110 AFIQPAMHNAVLPVFRITLPVLAPEVDHAPWRELEHDMKSEYAHLEKPEWKLGTGPPC 169
Qy 164 SWTFYHHLRPEILKEFTLHDHVRBEAQAFILGRV--NGSQPSTFVGVHVRGGDYVHVM 221
Db 170 SWTFHHLRQIRSEFTLHDHVRBEAQAFILGRV--NGSQPSTFVGVHVRGGDYVHVM 229
Qy 222 NWKGVVADRGYLEKALDMFRARYSSPVVYTSNGMAWCENINASRGDVVFAGNGIEGS 281
Db 230 KRWKGVVGDGYLQOADMWFRARYEAPVFFVYTSNGMEWCKENIDTSRGDVFAGNGREA 289
Qy 282 PAKDFALLTQCNTHTIMTIGTGFIAAYLAGDITVILANYTLPSDFPLKVKFKEAAFLPEW 341
Db 290 PARDFALLVQCNTHTIMTIGTGFIAAYLAGDITVILANYTLPSDFPLKVKFKEAAFLPEW 349
Qy 342 VGIADLSPL 351
Db 350 VGINADLSPL 359

```

RESULT 12

```

US-09-844-705-13
; Sequence 13, Application US/09844705
; Patent No. US2002013836A1
; GENERAL INFORMATION:
; APPLICANT: BOSWORTH, BRAD
; APPLICANT: VOGELI, PETER
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
; FILE REFERENCE: 21419/90368
; CURRENT APPLICATION NUMBER: US/09/844,705
; CURRENT FILING DATE: 2001-04-27

```



```
| : | : | : |
Db 328 VFLPPEGEPTK 339

RESULT 15
US-09-479-614-2
; Sequence 2, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-2

Query Match 5.5%; Score 113; DB 9; Length 496;
Best Local Similarity 20.58; Pred. No. 0.0061;
Matches 89; Conservative 40; Mismatches 127; Indels 176; Gaps 22;

QY 2 ASAQVFFSFPLA-----HFLIFVFVTSIIHLOQRIIVKLOPLSE 40
Db 69 ASIQAPLVFPLATCKGTATAPSVTLGCLVTGYFPPMPVITWDARSINKSV----- 121
QY 41 KELPMTQMSSG-----NTSEPFEMRRDSEQ----- 65
Db 122 -TLPATLQETSGLYTTTSHVTSGEWAQKFTCSVAHAESPTINKTVSACTMNFIPPTVK 180
QY 66 --HGNGELRGMTINSIGRLGNQMGVATLFAARMNGRLAFIPASMH-----NALAP 116
Db 181 LFHS-----SCNPLGDTG-----STQLCLISG---YVEGDMEVTLVDGQKATN 223
QY 117 IFRISLP-----VLHSDTAKKIPQNYHLNDWMBEERYHIFGHEVFTGYPCSWTFY 169
Db 224 IFPYTARGKQGVKVTSTHSEL-----NITQGEWVSQK-----TYICQVTY- 263
QY 169 HHLRPEILKEFTLHDHREAAQAFRLGRVNGSQPSFVGVHVRGDIYVHMPNWNKGVV 228
Db 264 -----OGFTFEDHARKCTESDPRGVSTYLSPPSPL-----DLYVHKSPKI-TCLV 307
QY 229 ADRGYLEKALDMFRARYSSPVFVVTSGM--AWCKENINASRGDVVFAGNGIEGSPAKDF 286
Db 308 VD-----LANTGMILTWREN-----GESVHPDP---M 333
QY 287 ALLPQCNHTIMTIGFTGIWAYLAGGDTIYLANIYLPDSP---FLKVFK-PEAAFLPE-W 341
Db 334 VKKQYNGTITVTSTPLVDANDWVEGET-YOCKVTHTDLPKDIVRSIAKAPGRFRFPPEVY 392
QY 342 VGIPADLSPLLK 353
Db 393 VFLPPEGEPTK 404
```

Search completed: May 27, 2003, 15:25:41
Job time : 19.4181 secs

B56392

beta-galactoside alphas,2-fucosyltransferase II - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: B56392
R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: B56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-354 <HIT>
A:CROSS-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357
C:Keywords: transmembrane protein

Query Match 64.9%; Score 1319.5; DB 2; Length 354;
Best Local Similarity 73.2%; Pred. No. 1.9e-102;
Matches 248; Conservative 27; Mismatches 47; Indels 17; Gaps 2;

QY 14 HFLFVFTSTIIHLQORIVKQLPSEKELPMTQMSSGNTSPSEMRDSEOHGNGELRG 73
DB 29 YLFTIFVSTVFHCHQRLAL-----VPAPWAYSARVVVP-----GHLPRG 71
QY 74 MFTINSIGRLNQMGGEYATLFALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKI 133
DB 72 MWTINAMGRIGNQMGGEYATLYALAKENGPRAPVPAQMHSTLAPIFRISLPVLHSDTASEV 131
QY 134 PQWNYHLNDWMEERYHHPGFHVRGTGTPCSTWTFYHHLRPELKEFTLHDHVRERAAQFL 193
DB 132 PQWNYHLNDWMEERYHHPGFHVRGTGTPCSTWTFYHHLRPELKEFTLHDHVRERAAQFL 191
QY 194 RGLRNGSQPSFTFVGHVHRRGDYVHMPNVKGVYADRGYLEKALDMPRARSPPVFTV 253
DB 192 RGLRNGSRPSTFVGHVHRRGDYVHMPNVKGVYADRGYLEKALDMPRARSPPVFTV 251
QY 254 SNGMAWCRENINASRGDVVYFAGNGTEGSPAKDFALLTQCNHFTMTIGTFGIWAAYLAGGD 313
DB 252 SNGMAWCRENINASRGDVVYFAGNGTEGSPAKDFALLTQCNHFTMTIGTFGIWAAYLAGGD 311
QY 314 TYILANYTLPSDFPKFKPEAAFLPEWGIADLSPLL 352
DB 312 TYILANYTAPDSFFHLVPKPEAAFLPEWGIATANNMGRAL 350

RESULT 3
A56392
beta-galactoside alphas,2-fucosyltransferase I - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: A56392
R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: A56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <HIT>
A:CROSS-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355
C:Keywords: transmembrane protein

Query Match 53.5%; Score 1086.5; DB 2; Length 373;
Best Local Similarity 57.6%; Pred. No. 5.7e-83;
Matches 209; Conservative 51; Mismatches 62; Indels 41; Gaps 6;

QY 16 LIFVFTV-----STIIHLQORIVK-----LQPLSKE-----LPMTTQMS----- 50
DB 11 LAFLVLCALSAFSLHLHLQDLFRNGLSLPCLERQVPAPVAIVCLPVTSPASNASSC 70
QY 51 SNTSPSEMRDSEOHGNGELRGMTINSIGRLNQMGGEYATLFALARMNGRLAFIPASM 110
DB 71 AGRPAAPS-----GIWTIHPDGRFGNQMGYATLLALAOULNGRRRPFILPAM 116

QY 111 HNALAPIFRISLPVLHSDTAKKIPWNYHLNDWMEERYHHPGFHVRGTGTPCSTWTFYH 170
DB 117 HAALAPVFRITLIPVLAPEVNRRTSKQLLLHDWMSSEYSRLDEDPLKFTGTFPCSWTFFHH 176
QY 171 LRPEILKEFTLHDHVRERAAQFLRLRV--NGSQSTFVGHVHRRGDYVHMPNVKGVY 228
DB 177 VREQIRREFTLHDHVRERAAQRLKRLGRTGARPTFVGHVHRRGDYVHMPNVKGVY 236
QY 229 ADRGYLEKALDMFRARYSSPVFVTSNGMAWCRENINASRGDVVYFAGNGIEGSPAKDFAL 288
DB 237 GDRAYLQOAMDWFRARHEAPFVVTSGMKWCWENIDASRGDVVYFAGNGLESSPAKDFAL 296
QY 289 LTQCNHTMTIGTFGIWAAYLAGGDYIYLYNTLPDSPLKVPKPEAAFLPEWGIADLS 348
DB 297 LTQCNHTMTIGTFGIWAAYLAGGDYIYLYNTLPDSPLKVPKPEAAFLPEWGIADLS 356
QY 349 SPL 351
DB 357 SPV 359
RESULT 4
A36047
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - human
C:Species: Homo sapiens (man)
C:Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 10-Sep-1997
C:Accession: A36047
R:Larsen, R.D.; Ernst, L.K.; Nair, R.P.; Lowe, J.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678, 1990
A:Title: Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-g
A:Reference number: A36047; MUID:90370848; PMID:2118655
A:Accession: A36047
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-365 <LAR>
A:CROSS-references: GB:M35531; NID:g183887; PID:g306830
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 52.0%; Score 1057.5; DB 2; Length 365;
Best Local Similarity 56.5%; Pred. No. 1.5e-80;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps 4;

QY 12 LAHFLIFVFTSTIIHLQO-----RIVKLQPLSEKELPMTQMSSGNTS 56
DB 11 LAFLVLCVLSVIFLHHDSPPHGLGSLPDRRLVTPPVAFCLCPGTAKGPNASSC 70
QY 57 PEMRRDSEOHGNGELRGMTINSIGRLNQMGGEYATLFALARMNGRLAFIPASMENALAP 116
DB 71 P-----QH-PASLSGTWTVYPNGRFGNQMGYATLLALAOULNGRRRPFILPAMHAALAP 122
QY 117 IFRISLPVLHSDTAKKIPWNYHLNDWMEERYHHPGFHVRGTGTPCSTWTFYHHLRPEIL 176
DB 123 VFRITLPVLAPEVDSRTPWRELQLDWMSSEYADLRDPFLKLSGFPSCSWTFFHHLREQIR 182
QY 177 KEFTLHDHVRERAAQFLRLRV--NGSQSTFVGHVHRRGDYVHMPNVKGVYADRGVY 234
DB 183 REFTLHDHVRERAAQVSLGQLRGLRGTDRPTFVGHVHRRGDYVHMPNVKGVYADRGVY 242
QY 235 EKALDMFRARYSSPVFVTSNGMAWCRENINASRGDVVYFAGNGIEGSPAKDFALLTQCNH 294
DB 243 ROAMDWFRARHEAPVFTVTSNGMEWCENIDFSQGDVTFAGDQGEATPKWDFALLTQCNH 302
QY 295 TMTIGTFGIWAAYLAGGDYIYLYNTLPDSPLKVPKPEAAFLPEWGIADLSPLKLA 354
DB 303 TMTIGTFGIWAAYLAGGDYIYLYNTLPDSPLKVPKPEAAFLPEWGIADLSPLKLA 362
QY 355 LTP 357
DB 363 AKP 365
RESULT 5

S46494
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) B - rat (fragment)
N:Alternate names: alpha-1,2-fucosyltransferase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 08-Oct-1999
C:Accession: S46494
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994.
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially expressed in the rat
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46494
A:Molecule type: mRNA
A:Residues: 1-159 <PI2>
A:Cross-references: EMBL:L26010; NID:9414816; PIDN:AAB41515.1; PID:g414817
C:Genetics: FTB
A:Gene: FTB
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 41.6%; Score 846; DB 2; Length 159;
Best Local Similarity 99.4%; Pred. No. 2.3e-63;
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 222 NWKGVADRGYLEKALDMFRARYSPFVVTSGMAWCRENINASRGDVFAGNGIEGS 281
DB 1 NWKGVADRGYLEKALDMFRARYSPFVVTSGMAWCRENINASRGDVFAGNGIEGS 60

QY 282 PAKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 341
DB 61 PAKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 120

QY 342 VGIPADLSPLKALTTPACPRSHFHLKAGVTCYVAGRAF 380
DB 121 VGIPADLSPLKALTTPACPRSHFHLKAGVTCYVAGRAF 159

RESULT 6
S51582
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1995 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C:Accession: S51582; S46493
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
submitted to the EMBL Data Library, November 1993
A:Reference number: S51582
A:Accession: S51582
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-142 <PIA>
A:Cross-references: EMBL:L26009; NID:9414814; PIDN:AAB41514.1; PID:g554438
Biochem. J. 300, 623-626, 1994
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially expressed in the rat
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46493
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 9-135 <PI2>
A:Cross-references: EMBL:L26009
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 28.6%; Score 581; DB 2; Length 142;
Best Local Similarity 73.2%; Pred. No. 2.6e-41;
Matches 104; Conservative 15; Mismatches 21; Indels 2; Gaps 1;

QY 162 PCSWTFYHHLRPEILKEFTLHDHVRERQAFLRLGVN--GSPSTFGVHVRGDIYVHV 219
DB 1 PCSWTFYHHLRPEILKEFTLHDHVRERQAFLRLGVN--GSPSTFGVHVRGDIYVHV 60

QY 220 MPNVKGVADRGYLEKALDMFRARYSPFVVTSGMAWCRENINASRGDVFAGNGIE 279
DB 61 MPNVKGVADRGYLEKALDMFRARYSPFVVTSGMAWCRENINASRGDVFAGNGIE 120

QY 280 GSPAKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 341
DB 121 GSPAKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 159

DB 121 GTPGKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 142
RESULT 7
T44328
hypothetical protein wbla [imported] - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
R:Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Gene 237, 321-332, 1999
A:Title: The genes responsible for O-antigen synthesis of Vibrio cholerae O139 are cl A:Reference number: Z27749; MUID:99453293; PMID:10521656
A:Accession: T44328
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-281 <YAM>
A:Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BAA33632.1; PID:g3721682
A:Experimental source: strain O22
C:Genetics:
A:Note: wbla
C:Superfamily: Vibrio cholerae hypothetical protein wbla

Query Match 10.6%; Score 214.5; DB 2; Length 281;
Best Local Similarity 26.5%; Pred. No. 2.5e-10;
Matches 79; Conservative 50; Mismatches 102; Indels 67; Gaps 14;

QY 74 MFTINSIGRLNQMGAYATLALARMNG---RL---AFIPASNMHNLALAFRISLPVLHS 127
DB 1 MIVMKISGGLGNLFQYAVGRAIAIOYGVPLKLDYSAYKNYKLHG---YLDQFNINA 56

QY 128 DTAKKIPWQNTL-----NOMWEERYHI-----PGHEVFRFT 159
DB 57 DIANED--EIFHLGSSNRLSLRRLGLWLNKNTYAEKQRTIYDVSVMQAPRY---LD.111

QY 160 GYPCSWTFYHHLRPEILKEFTLHDHVRERQAFLRLGVN--GSPSTFGVHVRGDIYVHV 219
DB 112 GYPCSWTFYHHLRPEILKEFTLHDHVRERQAFLRLGVN--GSPSTFGVHVRGDIYVHV 165

QY 220 MPNVKGVADRGYLEKALDMFRARYSPFVVTSGMAWCRENINASRGDVFAGNGIE 279
DB 166 HPEIG---VLDIDYKRAVDYIKETIAPVFFVSDVAVCKDNFIDSPVFIETDTE 222

QY 280 GSPAKDFALLTQCNTHTMTTGTGCIWAAYLAGDGTIYIANTLPDSPFLKVFKEAFLPEW 325
DB 223 ---IDLEMLMCCQCHNIVANSFSSWAAWLNLSNVDKIIVAPKTWMAENPKGVKWPDS 277

RESULT 8
H87911
protein B0205.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: H87911
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C- A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; A:Accession: H87911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <SFO>
A:Cross-references: GB:chr.I; PIDN:AAC16988.1; PID:g3150470; GSPDB:GN000019; CESP:B020 A:Note: contains weak similarity to fucosyltransferases
C:Genetics:
A:Gene: B0205.4
A:Map position: 1

Query Match 9.0%; Score 183; DB 2; Length 443;
Best Local Similarity 20.5%; Pred. No. 1.9e-07;


```
Matches 72; Conservative 53; Mismatches 128; Indels 96; Gaps 11;
QY 83 LGNOMGEYATLALARMGRIFAPA--SMHNA--LAFIFRISLPVLHSDWAKI-----133
Db 103 LGNLFQVAGLLSIARETGSGILLISTTTLRRAPDEFITENDSIQFVGEDLSRQAEOLN 162
QY 134 -----PWQNYHLNDMEERYRHIPGHVRFGTGPCSWTFYVHHLRPEILKE-FTLH 182
Db 163 ASKITLTSCCAYNRLSTILFNDRI-----IERIDGYFQNFYRPHPSQKIVKLFTEM 216
QY 183 DHVREAAQAF-----LRGLR 197
Db 217 DPVRKRVSNWLIYNIHPTNHRKPKESIVSIFSVFIPQLRVDFLENGLISLTVRNAR 276
QY 198 V-----NGSQ-----PSTFGVHVHRRGDYHVH--MPNVWKGWVADRGYLEKAL 238
Db 277 VIETNVANDQALEPEEDAFAKTMWGVHIRHGMDISMNSRNRHGHVDPPIEYKKRAI 336
QY 239 DMFRARYSSPVVVTNSGMACRENINASRGDVVPAGNIGSGSPAKDALLTQCHHTTMT 298
Db 337 QOISKIYENVAFIICSDNVAVARRNKLKGKTLHFF---CPGPREVDMAILKSCDSVIIS 393
QY 299 IGTFGIWAAYL---AGSDTIYLANVTLDPSPFLKVFKEPAEAFLEPWGVI 344
Db 394 TGTGQWWSAYLNVNASPDVYIYKHWPAPGSGVMEKTNKTEIFLKSWTAL 442

RESULT 9
T20745
hypothetical protein F11A5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20745
R:Gardner, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19319
A:Accession: T20745
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-363 <WILL>
A:Cross-references: EMBL:Z292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5
A:Experimental source: clone F11A5
C:Genetics:
A:Gene: CESP:F11A5.5
A:Map position: 5
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.4%; Score 170.5; DB 2; Length 363;
Best Local Similarity 23.4%; Pred. No. 1.7e-06;
Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;
QY 79 SIGRLGNOMGEYATLALARMGRILA--FIPASMH-----NALAFIFRISLPVLHSDT 129
Db 78 SSRGLGNHLFELASVLSIRELQRPVTFEENCYHEKMWEDSNTLIPGLMNHFLINGSV 137
QY 130 AKKIPWQNYH-----INDWMEERYRHIPGHVRFGTGPCSWTFYVHHLRPEILKEF 179
Db 138 PSSVKRVKFKQKCTDDPSLDLNDYEDYLHLTG-----THYQ-SWKYFSHMRNELI--- 188
QY 180 ILHDHVEEAQAFLGLRVNGSOPSTFVG-VHVHRRGDYVHVMPNVWKGV-VADRGYLEKA 237
Db 189 ---GYLKTENTY---MDLPKSGENTFITCVHVHRRGDLRV-----GFHVADENFIRSS 236
QY 238 LDMFRARYSSPVVVTNSGMACRENINASRGDVVPAGNIG-----VFEGDDYEFWDLSLRNPTSKINAFVSQ 278
Db 237 LNLISROVAKRANTAT-----VFEGDDYEFWDLSLRNPTSKINAFVSQ 278
QY 280 GSPAKDEALL-TCNHTIMTI--GTGFIWAAYLAGDGTIYLANVTLDPSPFLK----- 329
Db 279 NSPADLLLYAKSCNDVVLITAAHSTFCGWMGYFSKGNRVY---YT-DIQFTKDWILETG 333
QY 330 VFKEPAEAFLEPW 341
```

```
Db 334 EFISEDYILPHW 345
RESULT 10
T31916
hypothetical protein C17A2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T31916
R:Sammons, L.; Wohlmann, P.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C17A2.
A:Reference number: Z21098
A:Accession: T31916
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-348 <SAM>
A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00020; CESP:C17A2.4
A:Experimental source: strain Bristol N2; clone C17A2
C:Genetics:
A:Gene: CESP:C17A2.4
A:Map position: 2
A:Introns: 94/3; 133/3; 168/2; 272/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.7%; Score 156.5; DB 2; Length 348;
Best Local Similarity 20.9%; Pred. No. 2.3e-05;
Matches 72; Conservative 57; Mismatches 118; Indels 97; Gaps 18;
QY 12 LAHLFLIF-VFVTSIIHLQQRIVKLQPLSEKELPMTTOMSSGNTESPERRDSEQHNGE 70
Db 6 VVHSMKFRVFPYPAVLHNS-----TEHFLSSNLASP-----37
QY 71 LRGMFTINSIORLNGOMGEYATLALARMNGR--LAFIPASMHNALAFIFRISLP-----123
Db 38 -----SRLGNHIFEFASLGLSERLHRTPLFLVENEHFQKMLDTRKVPGLVEK 87
QY 124 --VLHSDTAKI---PWQ-----NVHLNDWMEERYRHIPGHVRFGTGPCSWTFYVHHL 171
Db 88 FTVINGSIFPSIKITPQKVCRRHENFEILEKIDKYLHLTGMEFYQ-----SNKYFPNM 141
QY 172 RPEILKKEFTLHDHVEEAQAFLGLRVNGSOPSTFVG-VHVHRRGDYVHVMPNVWKGV-VAD 230
Db 142 REQLL-----DFLDDSSQDF--GNLPRSNORTHVTCVHARRGDFVDV-----GFOAAD 187
QY 231 RGYLEKALDMFRARYSSPV-----FVYTSNGMACR---EN--INASRGDVVFRAGNG 277
Db 188 PDFIRNSVKYIAENFIPELEYKVKHKKVIFGDDLEFMRSILFNSVSTDEPEYMFPAEY 247
QY 278 I--EGSPAOKDFALLTQ--CNHTIMTI--GTGFIWAAYLAGDGTIY 316
Db 248 YISONSAPEDLTYSKQCDIVLISAPASTFGWIGYFSKGNKVF 291

RESULT 11
T32294
hypothetical protein K06H6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T32294
R:David, M.; Wohlmann, P.; Bauer, C.; Clarke, K.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid K06H6.
A:Reference number: Z21147
A:Accession: T32294
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-335 <DAV>
A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6
A:Experimental source: strain Bristol N2; clone K06H6.
C:Genetics:
A:Gene: CESP:K06H6.6
```

A;Map position: 5

A;Introns: 17/1; 32/2; 58/1; 111/3; 150/3; 184/2; 218/1; 243/3; 272/1; 303/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.6% Score 153.5; DB 2; Length 335;
Best Local Similarity 22.8%; Pred. No. 3.9e-05;
Matches 77; Conservative 50; Mismatches 118; Indels 93; Gaps 19;

QY 51 SGNTSEPMRDSQHGNGELGRGFTINSTGRGNQMGVATLALARMNGRLA--RIPA 108

Db 27 SATABEPLRQESFLIMKKRLSRMA--PTARLGNHMFELAAVLGISRLNRTATFFIED 84

QY 109 SMNALAPIFRISLP-----VLHSDTAKKIPW--QNYHLN-----DWMER 147

Db 85 EIYRMIESKEALPGLVGQFEILNG---KVPLYIKNTKLNRCCVFVDPLIHEHNDE 140

QY 148 YRHIPGHVFRGTGVCWIFVHHLRPEILKEFTLHDHVREBAQAFRLGRVNGSQP----- 203

Db 141 YLHLDGRFYQ-----AWKYPSPMRNELI-----GYLKTSENGSLPKSNE 180

QY 204 STFGV-VHVRGSDYVHVMNVKGVV-ADRGYLEKALDMF--RARYSSPVFVTSNG--- 256

Db 181 TSFVTCVHIRGDKRV-----GFAESDEWFIKAKEFVENKEATSKPCSHVVLFGDDL 234

QY 257 --MAWCREN-----INASRGDVVFAGNGIEGSPAKDFALLTQCNHTIMTI--GTGCIW 305

Db 235 PFMKNLYENESTHEVSNSPDDLLVYAKN-----NCDVLIITAPHSTFGW 280

QY 306 AAYLAGGDTIYLANI--TLPSDFFLK-VFKPRAAFLEPW 341

Db 281 MGYFSGDKGVYMDIRETRONVYRNGNLNPDYVYLPW 318

RESULT 12

T20572

hypothetical protein F08A8.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T20572

R:Harris, B.

submitted to the EMBL Data Library, October 1997

A;Reference number: Z19294

A;Accession: T20572

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-383 <WIL>

A;Cross-references: EMBL:Z99710; PIDN: CAB16868.1; GSPDB: GN00019; CESP: F08A8.5

A;Experimental source: clone F08A8

C:Genetics:

A;Gene: CESP:F08A8.5

A;Map position: 1

A;Introns: 16/1; 51/3; 77/1; 131/3; 170/3; 209/2; 244/1; 269/1; 297/3; 324/1; 354/3

C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.2% Score 146; DB 2; Length 383;

Best Local Similarity 21.4%; Pred. No. 0.0002;

Matches 87; Conservative 55; Mismatches 148; Indels 116; Gaps 16;

QY 13 AHFLIFVVTSTIHLQRIYKLOPLSEKLPMTTQMSGNTSEPMRDSQHGNGELR 72

Db 19 ANIYVITKITLSSHILEYFKQNSLSPKCEINNSQKSPNYEK-----TLQ 65

QY 73 GNFTINSTGRGNQMGVATLALARMNGRLAFTA-----SMNALAPIFRISLPVL 125

Db 66 LMLFAPSGGLGNKLFELISLHGATSLQKRAVINATNPSTFETLNRNIQPLF-----PKL 121

QY 126 HSD-TAKKIP--WQNYHLNDW-----MEERYRHIPGHVFRFTGYPCSWTFY 168

Db 122 ADQTLRIPIPSLVTHQQTNGRCVDDPDRFLNRSDQNLILDGHYFQ-----SFXY 175

QY 169 HHLRPEILKEFTLHDHVRE-EQAFLRLGRVNGSQPSTF-----VGVHVRGSDYVHVMNV 222

Db 176 HHIRPQ-----VREWLAPSKLQAMRAEILIPAKFRDDELICTHVRGDPQYDGLH 225

QY 223 VKGVVADRGYLEKALDMFRARYSSPVFVTSNGMWCRENINASRGDVVFAGNGIECP 282

Db 226 RPSDATTRAATDFLDLYLRKSHERVNVVLGN-----DIHFAYTVFEDRV 271

QY 283 AK-----DF-----ALLTQ-----CNHTIMTI--GTGFI 304

Db 272 AHFTFLOKPVNNSDYSLSPQISPSYTAILTFTLPEIDLAFSLFCVDVILITAFSSTFGW 331

QY 305 WAAYLAGGD-TIYLANITLPSDFLKFVKEPAAFLPEWVGIPADJS 349

Db 332 WLSYLAKRATATYVRDILLESKGVAGEMHPEDFYPEWIKLKTDLN 377

RESULT 13

T22068

hypothetical protein F41D3.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22068

R:White, S.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19508

A;Accession: T22068

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-500 <WIL>

A;Cross-references: EMBL:Z81537; PIDN: CAB04377.1; GSPDB: GN00019; CESP: F41D3.6

A;Experimental source: clone F41D3

C:Genetics:

A;Gene: CESP:F41D3.6

A;Map position: 1

A;Introns: 23/1; 59/1; 112/3; 150/3; 222/1; 279/1; 308/3; 326/3; 421/3

Query Match 6.8% Score 138.5; DB 2; Length 500;

Best Local Similarity 23.4%; Pred. No. 0.0012;

Matches 74; Conservative 48; Mismatches 105; Indels 89; Gaps 16;

QY 74 MFTNSTIGRLNOMGEYATLALARMNGRLAPIPASMNALAPIR----- 119

Db 49 MTIVYKGLGNQLFEVLISLLGIARKKRIA-----VFNSSDPVLQSLNLEFLNQKLPRIS 103

QY 120 --ISLPVLHSDTAK-----KIPWONYHLNDWMEERYRHIPGHVFRFTGYPCSWTFY 168

Db 104 EOVISVPIEPSETTFRATSSCCRYELSDNLHAD--ESKFLVIEGHYFQ-----SYKIF 155

QY 169 HHLRPEILKEFTLHDHVREBAQAFRLGRVNGSQPSTFVGVHVRGSDYV---HYVMPNVW 224

Db 156 ADMKLSI-KEWL---KPEDPEKFRMISKTESQRHK-TCVHVRGDELTDQOH----- 203

QY 225 KGVVADRGYLEKALDMFRARYSSPVFVTSNGMWCRENINASRGDVVFAGNGIEGSPAK 284

Db 204 --AGTDSNYTISAIHLRLSLXGVIFIM-SNDPKVWKVHI-ADHLDY-----QK 248

QY 285 DFALL-----TQCNTIMTI--GTGFWAAVLAGGDT-IYLANI-TLPDS 325

Db 249 DIRIMKTMEDAIDDLHFSQYCDSVLITAFSSTFGWIGLSKNQSAVYVRDIRETKDQ 308

QY 326 PFLKVPKPEAAFLPEW 341

Db 309 VOLQMTK-EDFYPTW 323

RESULT 14

H71976

probable alpha(1,2)fucosyltransferase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A;Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000

C:Accession: H71976

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999


```
FT VARIANT 138 138 R -> C (IN XHOSA POPULATION).
FT FTID=VAR_003423.
FT VARIANT 140 140 I -> F (IN JAPANESE SEJ ALLELE; NON-
FT SECRETOR).
FT FTID=VAR_003424.
FT VARIANT 172 172 D -> N (IN XHOSA POPULATION).
FT FTID=VAR_003425.
SQ SEQUENCE 343 AA; 39017 MW; 12066D9CF175E13A CRC64;

Query Match 73.1%; Score 1485.5; DB 1; Length 343;
Best Local Similarity 77.9%; Pred. No. 3.4e-118;
Matches 275; Conservative 30; Mismatches 37; Indels 11; Gaps 2;

QY 1 MASQVPPSPPLAHFLIFVFTSTIIHQQRIVKLQPLSEKELPMTTQMSSGNTESP60
D 1 MLVQMPFSPMAHFLIFVFTSTIIHQQRIVKLQPLSEKELPMTTQMSSGNTESP60
QY 61 RDSQHGNGELRGMTINSIGRLQNGMEYATLALARMNGRLAFIPASMHNALAPIFRI 120
D 50 STKALGPSQLRGMTINAIIGRLQNGMEYATLALAKMNGRAAFIPQMHSTLAPIFRI 109
QY 121 SLPLVHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPGCSWTFYHHLRPEILKEFT 180
D 110 TLPVLHSATSRIPWNYHLNDWMEERYRHIPGHFVFTGYPGCSWTFYHHLRPEILKEFT 169
QY 181 LHDHVRERAAQFLRLGLRVNGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
D 170 LHDHVRERAAQFLRLGLRVNGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 229
QY 241 FRARYSPFVFTVNSNGMAWCRENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIG 300
D 230 FRARYSSILFVFTVNSNGMAWCRENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIG 289
QY 301 TFGIWAAYLAGGDTIYLYANTLPDPSPLKFKPEAAFLPEWVGIPADLSPLK 353
D 290 TFGIWAAYLAGGDTIYLYANTLPDPSPLKFKPEAAFLPEWVGIPADLSPLK 342

RESULT 3
FUT2_BOVIN STANDARD; PRT; 344 AA.
AC Q28113;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (alpha(1,2)FT 2)
DE (Fucosyltransferase 2).
GN FUT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Petit J.M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X99620; CAA67931.1; -.
```

```
DR InterPro: IPR002516; GI 11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Golgi stack; Signal-anchor.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 29 344 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 344 AA; 39320 MW; F35AC33F1B7B9F68 CRC64;

Query Match 72.1%; Score 1466; DB 1; Length 344;
Best Local Similarity 77.1%; Pred. No. 1.5e-116;
Matches 272; Conservative 27; Mismatches 44; Indels 10; Gaps 1;

QY 1 MASQVPPSPPLAHFLIFVFTSTIIHQQRIVKLQPLSEKELPMTTQMSSGNTESP60
D 1 MFSQCTFEFFPTAFILFVFTSTIIHQQRIVKLQPLSEKELPMTTQMSSGNTESP60
QY 61 RDSQHGNGELRGMTINSIGRLQNGMEYATLALARMNGRLAFIPASMHNALAPIFRI 120
D 51 TPRSPPQRPQLKGMWTINAIIGRLQNGMEYATLALAKMNGRAAFIPQMHSTLAPIFRI 110
QY 121 SLPLVHSDTAKKIPWNYHLNDWMEERYRHIPGHFVFTGYPGCSWTFYHHLRPEILKEFT 180
D 111 TLPVLHDATAKSPWQNYHLNDWMEERYRHIPGHFVFTGYPGCSWTFYHHLRPEILKEFT 170
QY 181 LHDHVRERAAQFLRLGLRVNGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
D 171 LHAHVRERAAQFLRLGLRVNGSQPSTFGVHVRRGDYVHVMPNVKGVVADRGYLEKALDM 230
QY 241 FRARYSPFVFTVNSNGMAWCRENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIG 300
D 231 FRARYSAPFVFTVNSNGMAWCRENINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIG 290
QY 301 TFGIWAAYLAGGDTIYLYANTLPDPSPLKFKPEAAFLPEWVGIPADLSPLK 353
D 291 TFGIWAAYLAGGDTIYLYANTLPDPSPLKFKPEAAFLPEWVGIPADLSPLK 343

RESULT 4
FUT2_MOUSE STANDARD; PRT; 368 AA.
AC P97333;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (alpha(1,2)FT 2)
DE (Fucosyltransferase 2) (FUT-III).
GN FUT2 OR SECL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RC STRAIN=ICR;
RT "Molecular cloning and expression of a mouse GDP-L-fucose: beta-D-
RT galactoside 2-alpha-L-fucosyltransferase.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Lin B., Hayashi Y., Saito M., Sakakihara Y., Yandigisawa M.,
RA Tamori M.;
RT "Molecular cloning and expression of a GDP-L-fucose:beta-D-galactoside
RT 2-alpha-L-fucosyltransferase in murine gastrointestinal tract.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
```

```

CC CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC CC -!- PATHWAY: Glycosylation.
CC CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SEC1) WHICH ENCODE
CC CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: Y09882; CAA71008.1; -.
CC CC EMBL: AF113532; AAD25351.1; -.
CC CC MGD: MGI:109374; Fut2
CC CC InterPro: IPR002516; GT_11.
CC CC Pfam: PF01531; Glyco_transf_11; 1.
CC CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC CC Signal-anchor; Golgi stack.
CC CC DOMAIN 1 20
CC CC TRANSMEM 21 41
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC LUMENAL, CATALYTIC (POTENTIAL).
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 195 195
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 289 289
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 315 315
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC SEQUENCE 368 AA; 41464 MW; 4093853BB37303B CRC64;
CC CC -----
CC CC Query Match 67.1%; Score 1363.5; DB 1; Length 368;
CC CC Best Local Similarity 75.9%; Pred. No. 7.5e-108;
CC CC Matches 258; Conservative 24; Mismatches 39; Indels 19; Gaps 3;
CC CC -----
Qy 11 PLAHPLIF--VFVSTIHLQORIVKLPSEKELPMTQMSGNTSEPMRRDSEQHGN 68
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 PLSTYLFVFVSTVSTFCHRRLL-----GLVPAPWASPSLVVFPFRHP 66
Qy 69 GELRGMTINSTIGRLGNMGGEYATLAFALARMNGRLAFIPASHMNAALAFIRLSPLVLHSD 128
Db ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 RE--GMFTIRVKGRGLGNMGGEYATLAFALARMNGRLAFIPASHMNAALAFIRLSPLVLHSD 124
Qy 129 TAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDVREE 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
125 TAKRIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDVREE 184
Qy 189 AQAFRLGLRVNGSQPSTFVGHVHVRGDDYVHVPYKVGWVADRGYLEKALDMFRARYSSP 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 AQAFRLGLRVNGSQPSTFVGHVHVRGDDYVHVPYKVGWVADRGYLEKALDMFRARYSSP 244
Qy 249 VFVVTNSGMWMCNENINASRGDGVVFNAGNGIEGSPAKOFALLTCQNTHTMTGTGFIWAAY 308
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
245 VFVVTSDDMWCKRSITASRGDGVAFAGNGLOGSPAKDIALMQCNHTVITLTGFIWAAY 304
Qy 309 LAGGDTIYLANITLPDPSFLKVKFEPAFLPEWVGIPADL 348
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 LTGGDTVYLANFTQPSFHTVFKPEAAFLPEWVGIAADL 344
CC CC -----
CC CC RESULT 5
CC CC FUT2_RABIT
CC CC ID FUT2_RABIT STANDARD; PRT; 354 AA.
CC CC AC Q10983;
CC CC DT 01-OCT-1996 (Rel. 34, Created)
CC CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC CC DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood
CC CC group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-
CC CC galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)
CC CC (Fucosyltransferase 2).

```

```

GN FUT2 OR SEC1 OR RFT-II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.
RT "Molecular cloning and expression of two types of rabbit beta-
RT galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: X80225; CAA56512.1; -.
CC CC InterPro: IPR002516; GT_11.
CC CC Pfam: PF01531; Glyco_transf_11; 1.
CC CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC CC Signal-anchor; Golgi stack.
CC CC DOMAIN 1 22
CC CC TRANSMEM 23 43
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC LUMENAL, CATALYTIC (POTENTIAL).
CC CC CARBOHYD 197 197
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 291 291
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 317 317
CC CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6CCB4 CRC64;
CC CC -----
CC CC Query Match 64.9%; Score 1319.5; DB 1; Length 354;
CC CC Best Local Similarity 73.2%; Pred. No. 3.7e-104;
CC CC Matches 248; Conservative 27; Mismatches 47; Indels 17; Gaps 2;
CC CC -----
Qy 14 HFLIPFVFTSTIIHQQRIVKLPSEKELPMTQMSGNTSEPMRRDSEQHNGELRG 73
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 YFLETFIVYSTVFHCHORLAL-----VPAPWAYASRVVWVP-----GHLPRG 71
Qy 74 MFTINSIGRLGNMGGEYATLAFALARMNGRLAFIPASHMNAALAFIRLSPLVLHSDTAKKI 133
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72 MWTINAMGRGLGNMGGEYATLAKENGRPAYIPQMSTLAPIRISLPVLHSDTASRV 131
Qy 134 PQQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDVREAAQFL 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 PQQNYHLNDWMEERYRHIPGVVRLTGYPCSWTFYHHLRPEILKEFTLHDVREAAQFL 191
Qy 194 RGLRVNGSQPSTFVGHVHVRGDDYVHVPYKVGWVADRGYLEKALDMFRARYSSPVVFT 253
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
192 RGLRVNGSRPSTFVGHVHVRGDDYVHVPYKVGWVADRGYLEKALDMFRAPTPPVVFT 251
Qy 254 SNGMWCNENINASRGDGVVFNAGNGIEGSPAKDFALLTCQNTHTMTGTGFIWAAYLAGCD 313
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db 252 SNGMAWCRENDASRGDVFAGNGLEGSPAKDFALLQCNHTVMTICTGFGWAAVLTGGD 311
QY 314 TIVLANTYLPDSPLKVKFKEPAALPEWVGIPADLSPLL 352
Db 312 TVYLANATPDSPLKVKFKEPAALPEWVGITANKGRAL 350

RESULT 6
FUT1_MOUSE STANDARD; PRT; 376 AA.
ID FUT1_MOUSE
AC O09160;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)Ft 1)
DE (Fucosyltransferase 1).
GN FUT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTH Swiss;
RX MEDLINE=97454449; PubMed=9355741;
RA Domino S.E., Hiraiwa N., Lowe J.B.;
RT "Molecular cloning, chromosomal assignment and tissue-specific
RT thymic and epididymal epithelial cells.";
RL Biochem. J. 327:105-115(1997).
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- TISSUE SPECIFICITY: IN THE ADULT, HIGHLY EXPRESSED IN PANCREAS,
CC TESTIS AND EPIDIDYMIS AND TO A LESSER EXTENT IN THYMUS, LONG,
CC STOMACH, SMALL INTESTINE, COLON, SPLEEN AND UTERUS. NOT EXPRESSED
CC IN BRAIN, HEART, SKELETAL MUSCLE, KIDNEY, LIVER AND BONE MARROW.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U90553; AAC53492.1; -
CC DR MGD; MG1:109375; Fut1.
CC DR InterPro: IPR002516; Gt_11.
CC DR Pfam: PF01531; Glyco_transf_11; 1.
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
CC FT TRANSF 9 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC FT DOMAIN 27 376 LUMENAL, CATALYTIC (POTENTIAL).
CC FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 376 AA; 42255 MW; 21FD24CFE204106 CRC64;

Query Match 54.5%; Score 1107; DB 1; Length 376;
Best Local Similarity 60.5%; Pred. No. 3.5e-86;
Matches 210; Conservative 44; Mismatches 83; Indels 10; Gaps 3;

QY 21 VTSLIHLQORIVKQLPSKELPMVTMTQSSNGTSEPMRDSQHGNGELRGFTTNSI 80
Db 38 LTUSVLCSYHLK-SPVAMVCLPPLPQTSGNSPCEQ-----SSLSGTWITTPG 88

```

```

QY 81 GRLGNMGGEYATLALARMNGRLAPIPASMNALAPIPRISLPVLHSDTAKKIPWNIHL 140
Db 89 GRFGNMGQYATLLALAQNLGRQAFIQEPMEHAALAPVFRISLPVLDPDVDSLTPWQHLVL 148
QY 141 NDWMEERYRHIPGPHVFRFTGYPCSWTFYVHHLRPELTKEFTLHDHVRBEAQAFRLGLRVNG 200
Db 149 HDWMESEYSHLSDPFLKLSGPPCSWTFFHHLRBEQRRREFTLHNLHREGAQYLLSGLRGP 208
QY 201 SQPS-IFVGVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAW 259
Db 209 ASPAHTFVGVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAW 268
QY 260 CREMINASRGDVFVAGNGIEGSPAKDFALLTQCQHTTINTICTGFIWAAVLAGGDIILAN 319
Db 269 CLENIIDTSHGDDVFVAGNGIEGSPAKDFALLTQCQHTTINTICTGFIWAAVLAGGDIILAN 328
QY 320 YTLPSDFLKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFL 366
Db 329 FTLPSDFLKVFKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFL 375

```

RESULT 7

```

FUT1_MOUSE STANDARD; PRT; 376 AA.
ID FUT1_MOUSE
AC O10980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)Ft 1)
DE (Fucosyltransferase 1).
GN FUT1 OR FTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Colon cancer;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-
RT fucosyltransferase genes.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 169-310 FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=94280382; PubMed=8010942;
RA Piau J.-P., Labarrière N., Dabouis G., Denis M.G.;
RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
RT differentially expressed throughout the rat colon.";
RL Biochem. J. 300:623-626(1994).
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB015637; BAA31130.1; -
CC EMBL; AB006137; BAA21741.1; -
CC DR

```


OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370848; PubMed=2118655;
RA Larsen R.D., Ernst L.K., Nair R.P., Lowe J.B.;
RA "Molecular cloning, sequence, and expression of a human GDP-L-
RT fucose-beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can
RT form the H blood group antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6674-6678(1990).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT ALLELES.
RX MEDLINE=97240210; PubMed=9122901;
RA Wagner F.F., Flegel W.A.;
RA "Polymorphism of the h allele and the population frequency of
RT sporadic nonfunctional alleles.";
RL Transfusion 37:284-290(1997).
RN [3]
RP VARIANT HTS-164.
RX MEDLINE=94286534; PubMed=7912436;
RA Kelly R.J., Ernst L.K., Larsen R.D., Bryant J.G., Robinson J.S.,
RA Lowe J.B.;
RA "Molecular basis for H blood group deficiency in Bombay (Oh) and
RT para-Bombay individuals.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5843-5847(1994).
RN [4]
RP VARIANT ARG-242.
RX MEDLINE=97445117; PubMed=92399444;
RA Koda Y., Soejima M., Johnson P.H., Smart E., Kimura H.;
RA "Missense mutation of FUT1 and deletion of FUT2 are responsible for
RT Indian Bombay phenotype at ABO blood group system.";
RL Biochem. Biophys. Res. Commun. 238:21-25(1997).
RN [5]
RP FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
RT ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
RL SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
RT SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
RL SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
RN [6]
RP CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
RT alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
RN [7]
RP PATHWAY: Glycosylation.
RN [8]
RP SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
RT FORM IN TRANS CISTERNAE OF GOLGI.
RN [9]
RP POLYMORPHISM: NONFUNCTIONAL MUTANT OF FUT1 ARE THE CAUSE OF THE H-
RT BOMBAY BLOOD GROUP.
RN [10]
RP MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
RT GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
RL TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
RT MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
RN [11]
RP SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
RN [12]
RP This SWISS-PROT entry is copyrighted. It is produced through a collaboration
RT between the Swiss Institute of Bioinformatics and the EMBL outstation -
RL the European Bioinformatics Institute. There are no restrictions on its
RT use by non-profit institutions as long as its content is in no way
RL modified and this statement is not removed. Usage by and for commercial
RT entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RL or send an email to license@isb-sib.ch).
RN [13]
RP EMBL: M35531; AAA52639.1; -
RT EMBL: Z59587; CA93435.1; -
RL PIR: A36047; A36047.
RN [14]
RP Genew: HGNC:4012; FUT1.
RN [15]
RP MIM: 211100; -
RN [16]
RP InterPro: IPR002516; GT_11.
RN [17]
RP Pfam: PF01531; Glyco.transf.11; 1.
RN [18]
RP Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
RT Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.
RL DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
RN [19]
RP TRANSMEM 9 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
RT DOMAIN 26 365 LUMENAL, CATALYTIC (POTENTIAL).
RN [20]
RP CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
RT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
RN [21]
RP VARIANT 154 Y -> C (IN BOMBAY H-)

FT VARIANT 164 164 /FTid=VAR_003417.
FT L -> H (IN PARA-BOMBAY).
FT /FTid=VAR_009708.
FT VARIANT 171 171 W -> C (IN BOMBAY H-).
FT /FTid=VAR_003418.
FT VARIANT 242 242 L -> R (IN BOMBAY H-).
FT /FTid=VAR_009709.
FT VARIANT 259 259 V -> E (IN BOMBAY H-).
FT /FTid=VAR_003419.
FT VARIANT 315 315 A -> V (IN BOMBAY H-).
FT /FTid=VAR_003420.
FT VARIANT 349 349 W -> C (IN BOMBAY H-).
FT /FTid=VAR_003421.
SQ SEQUENCE 365 AA; 41251 MW; 4F442EC375C9D9E CRC64;
Query Match 52.0%; Score 1057.5; DB 1; Length 365;
Best Local Similarity 56.5%; Pred. No. 5.4e-82;
Matches 205; Conservative 46; Mismatches 87; Indels 25; Gaps
QY 12 LAHLFLIFVFTSTIIHQQ-----RIVKLQPLSEKELPMTTCMSSGNTES 56
Db 11 LAFLVLCVLSVIFFLHIHODSPHGLSLILCPDRRLVTPPVVAILCPTANGPNASSC 70
QY 57 PEMRRDSEQHNGELRGMPITINSIGRLGNQMGAYATLALARMNGRLAFIPASMMNALAP 116
Db 71 P-----QH-PASLSGTWTVPNGRFGNQMGQYATLLALQNGRRRAFILPAMHAALAP 122
QY 117 IFRISIPVLHSDTAKKIPWNYHLNDWMEEYRHLPFHVFRVETGYPCSWTFVYHLRPEIL 176
Db 123 VERITLPLVAPEDVSTPWRLEQLHDMSEYADLRDPFLKUSFPFCSTWTFPHHLEIR 182
QY 177 KEFTLHDHVRBEAQAFRLGRV--NGSQPSTFVGVHVRGDIYVHVMPNWKGVADRGYL 234
Db 183 REFTLHDHLEEAQSVGLQRLGRGDRPRTFVGVHVRGDIYVHVMPNWKGVADRGYL 242
QY 235 EKALDMFRARYSSPVFVTSNGMACRENINASRGVVFAGNIGSPAKDPALLTQCNH 294
Db 243 RQAMDWFRARHEAPVFTVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNH 302
QY 295 TMTTGTGFGWYAYLAGDITIYLANVTPDSPFLKVFKEPEAFLEPWGIPADLSPLLKA 354
Db 303 TMTTGTGFGWYAYLAGDITIYLANVTPDSPFLKVFKEPEAFLEPWGIPADLSPLLKA 362
QY 355 LTP 357
Db 363 AKP 365
RESULT 10
FUT1_PIG STANDARD; PRT; 365 AA.
ID FUT1_PIG
AC Q29043; Q19101;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)F 1)
DE (Fucosyltransferase 1).
GN FUT1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96217559; PubMed=8613146;
RA Conney S., Moutouris E., McKenzie I.F., Sandrin M.S.;
RA "Molecular cloning of the gene coding for pig alpha1-->2
RT fucosyltransferase.";
RL Immunogenetics 44:76-79(1996).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS THR-103 AND GLN-286.
RX MEDLINE=97468270; PubMed=9321466;

Query Match	4.5%;	Score 91;	DB 1;	Length 442;
Best Local Similarity	22.7%;	Pred. No. 3.2;		

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:52:27 ; Search time 45.4049 Seconds
(without alignments)
1050.630 Million cell updates/sec

Title: US-10-040-863-8_COPY_23_380

Perfect score: 1921

Sequence: 1 STIIHLQQRIVKLPLSEKE.....RSHFLKAKGVTCYVAGRAF 358

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	380	21	Rat hepatoma H35 C
2	1921	100.0	380	23	Rat hepatoma H35 C
3	1896	98.7	353	21	Rat hepatoma H35 c
4	1896	98.7	353	23	Rat hepatoma H35 c
5	1451	75.5	340	19	Porcine secretor t
6	1451	75.5	340	19	Pig secretor. Sus
7	1407.5	73.3	343	19	Human Sec2 protein
8	1397	72.7	344	21	Human Sec2 catalyt
9	1051.5	54.7	365	12	Human GRB-2. Homo
10	1051.5	54.7	365	16	GDP-L-fucose-beta-

11	1051.5	54.7	365	16	2-Alpha-fucosyltra
12	1051.5	54.7	365	16	2-Alpha-fucosyltra
13	1051.5	54.7	365	17	Human H-transferas
14	1051.5	54.7	365	18	Human alpha 1,2 fu
15	1051.5	54.7	365	18	Human alpha(1,2)-f
16	1051.5	54.7	365	21	Human H-transferas
17	1046.5	54.5	365	19	Pig H transferase.
18	1046.5	54.5	365	19	Pig alpha-1-2 fucco
19	1046.5	54.5	365	23	Swine alpha(1,2)
20	1044.5	54.4	365	20	Swine alpha(1,2) f
21	1044.5	54.4	365	20	Swine alpha(1,2) f
22	1043.5	54.3	365	15	A glycosyltransfer
23	990	51.5	357	20	X. laevis alpha-1,
24	288.5	15.0	110	21	Human secreted pro
25	141	7.3	300	21	Human secreted pro
26	141	7.3	300	22	Helicobacter pylor
27	139.5	7.3	287	23	Bacteroides fragil
28	103	5.4	661	22	Propionibacterium
29	101	5.3	690	23	Human protein cont
30	97	5.0	317	13	Growth Factor Rece
31	96.5	5.0	355	21	C. elegans alpha-1
32	95	4.9	1515	22	Novel human diagno
33	93	4.8	1073	22	Novel human diagno
34	92	4.8	984	22	Human polypeptide,
35	89.5	4.7	586	21	Amino acid sequenc
36	89.5	4.7	604	22	Propionibacterium
37	89	4.6	353	16	Porphyromonas geng
38	89	4.6	394	22	Human protein sequ
39	89	4.6	1031	22	Novel signal trans
40	89	4.6	2057	22	Angiogenesis assoc
41	89	4.6	2058	23	Novel human protei
42	88.5	4.6	1416	22	Novel human secret
43	88	4.6	217	23	Mouse ischaemic co
44	88	4.6	629	22	Novel human diagno
45	87	4.5	217	16	Human GRB-2. Homo

ALIGNMENTS

RESULT 1

AA36104

ID AA36104 standard; Protein; 380 AA.

XX

AC AA36104;

XX

DT 19-FEB-2001 (first entry)

XX

DE Rat hepatoma H35 cell alpha1-2fucosyltransferase.

XX

KW Rat: alpha1-2fucosyltransferase; cytostatic; neuroprotective;

KW neotropic; gene therapy; Fucalalpha1-2galbeta1-3galNAc; immunotherapy;

KW immunosuppression; cancer; neurological disease;

XX small cell lung carcinoma.

XX Rattus norvegicus.

XX

PN WO2000064464-A1.

XX

PD 02-NOV-2000.

XX

PF 23-APR-1999; 99WO-US07384.

XX

PR 23-APR-1999; 99WO-US07384.

XX

PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.

PI

PI Holmes EH, Sherwood AL;

XX

XX WPI; 2000-687262/67.

DR

DR N-PSDB; AAC67965.

XX

PT New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for

PT preparation of fucosyl GM1 which is useful as a nutritional composition
PT or immunotherapeutic for cancer and neurological diseases -
XX
PS Claim 1; Fig 5; 91pp; English.
XX
CC The present sequence was given in a specification relating to an isolated
CC rat ganglioside GM1-specific algal-2fucosyltransferase protein. The
CC protein or its cellular fraction is useful for synthesis of a molecule
CC comprising Fucal-2Gal-3GalNAc, a glycolipid, glycoprotein,
CC glycolipoprotein or a free oligosaccharide comprising
CC Fucal-2Gal-3GalNAc. The method involves contacting
CC algal-2fucosyltransferase with GDP-fucose and a molecule or glycolipid,
CC glycoprotein, glycolipoprotein or oligosaccharide having a terminal
CC Gal-3GalNAc group. It is also useful for synthesis of fucosyl-GM1
CC by contacting the protein with GDP-fucose and ganglioside GM1. The
CC obtained glycoproteins, glycolipoproteins, glycolipids and
CC oligosaccharides are useful as nutritional compositions and fucosyl-GM1
CC is useful for inducing an immunotherapeutic or immunosuppressive action
CC against cancer, neurological disease or small cell lung carcinoma.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1921; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.7e-201;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STIIHQQRIVKLOPLSEKELPMTTQSSGNTESPMRRDSEQHNGELRGMTINSIGR 60
Db STIIHQQRIVKLOPLSEKELPMTTQSSGNTESPMRRDSEQHNGELRGMTINSIGR 82
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPQWNYHLND 120
Db LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPQWNYHLND 142
QY 121 WMEERYRHPGHVFTGYPGCSWTYHHLRPEILKEFTLHDHVRSEAQAFLRGLRVNGSQ 180
Db WMEERYRHPGHVFTGYPGCSWTYHHLRPEILKEFTLHDHVRSEAQAFLRGLRVNGSQ 202
QY 181 PSTFVGVRGSDYVHVMPNVWKGVDGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db PSTFVGVRGSDYVHVMPNVWKGVDGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNTIMTIGTGIWAAYLAGGDTIYLANVTL 300
Db NINASRGDVFVAGNIEGSPAKDFALLTQCNTIMTIGTGIWAAYLAGGDTIYLANVTL 322
QY 301 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAGVTCVYAGRAF 358
Db PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAGVTCVYAGRAF 380

RESULT 2
AAEI6622 standard; Protein; 380 AA.
XX AAEI6622;
XX
XX
DT 09-APR-2002 (first entry)
XX
XX Rat hepatoma H35 cell algal-2Fuct.
XX

KW Rat; algal-2fucosyltransferase; algal-2Fuct; antisense therapy;
KW galactose beta1-3N-acetylglactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytosolic; ganglioside; GM1;
KW cell transformation.
XX

OS Rattus norvegicus.
XX
XX Key Location/Qualifiers
FH Domain 1..27
FT Modified-site /note= "Intracellular/Transmembrane domain"
FT Modified-site 199

FT Region /note= "N-glycosylated"
FT 221..380
FT Modified-site /note= "Region which overlaps rat FTH"
FT 265
FT Modified-site /note= "N-glycosylated"
FT 293
FT Modified-site /note= "N-glycosylated"
PN US6329170-B1.
XX
PD 11-DEC-2001.
XX
PF 23-APR-1999; 99US-0298886.
XX
PR 23-APR-1999; 99US-0298886.
PA (NHO-) NORTHWEST HOSPITAL.
XX
PI Holmes EH, Sherwood AL;
XX
XX WPI: 2002-121132/16.
DR N-PSDB; AAD27207.
XX
PT Rat hepatoma H35 cell algal-2fucosyltransferase, useful for producing
PT GM1-specific algal-2fucosyltransferase enzyme by recombinant
PT techniques and for detecting oncogenic transformation of test tissues -
XX
PS Claim 3; Fig 5; 41pp; English.
XX
CC The invention relates to rat GM1-specific algal-2fucosyltransferase
CC (algal-2Fuct) enzyme and its corresponding nucleic acid. This nucleic
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,
CC a terminal galactose beta1-3N-acetylglactosamine (Galbeta1-3GalNAc),
CC saccharide. Algal-2Fuct DNA is useful for producing rat algal-2Fuct
CC protein by recombinant techniques. Algal-2Fuct DNA is useful for the
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.
CC Algal-2Fuct DNA is useful for detecting oncogenic transformation which
CC involves assaying for changes in expression of algal-2 Fuct. Since
CC algal-2Fuct is activated in cell transformation, antisense sequences
CC derived from algal-2Fuct DNA are useful for inhibiting, suppressing
CC or treating cancer. Algal-2Fuct DNA is useful in gene therapy and
CC antisense therapy. The present sequence is rat hepatoma H35 cell
CC algal-2Fuct.
SQ Sequence 380 AA;

Query Match 100.0%; Score 1921; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.7e-201;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STIIHQQRIVKLOPLSEKELPMTTQSSGNTESPMRRDSEQHNGELRGMTINSIGR 60
Db STIIHQQRIVKLOPLSEKELPMTTQSSGNTESPMRRDSEQHNGELRGMTINSIGR 82
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPQWNYHLND 120
Db LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPQWNYHLND 142
QY 121 WMEERYRHPGHVFTGYPGCSWTYHHLRPEILKEFTLHDHVRSEAQAFLRGLRVNGSQ 180
Db WMEERYRHPGHVFTGYPGCSWTYHHLRPEILKEFTLHDHVRSEAQAFLRGLRVNGSQ 202
QY 181 PSTFVGVRGSDYVHVMPNVWKGVDGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db PSTFVGVRGSDYVHVMPNVWKGVDGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNTIMTIGTGIWAAYLAGGDTIYLANVTL 300
Db NINASRGDVFVAGNIEGSPAKDFALLTQCNTIMTIGTGIWAAYLAGGDTIYLANVTL 322
QY 301 PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAGVTCVYAGRAF 358
Db PDSPLKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAGVTCVYAGRAF 380

	Query Match	98.7%;	Score 1896;	DB 21;	Length 353;	
	Best Local Similarity	100.0%;	Pred. No. 1.3e-198;			
	Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	6	QQQIVKQLQPLSEKELPWTQMS	SGNTSPESPMRRDSEQHG	NGELRGMTINSIGRLGNQM	65	
Db	1	QQQIVKQLQPLSEKELPWTQMS	SGNTSPESPMRRDSEQHG	NGELRGMTINSIGRLGNQM	60	
QY	66	GEYATLFAALRMNGRLAFIPAS	MHNALAPIFRISLPVLHSDT	TAKKIPQWNYHLNDWMEER	125	
Db	61	GEYATLFAALRMNGRLAFIPAS	MHNALAPIFRISLPVLHSDT	TAKKIPQWNYHLNDWMEER	120	
QY	126	YRHPGHEVFTGYPCSWTYTHL	RPEILKEFTLHDHVREAAQFL	RLGRVNGSQPSFTV	185	
Db	121	YRHPGHEVFTGYPCSWTYTHL	RPEILKEFTLHDHVREAAQFL	RLGRVNGSQPSFTV	180	

CC Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which
 CC involves assaying for changes in expression of alpha1-2 Fuct. Since
 CC alpha1-2Fuct is activated in cell transformation, antisense sequences
 CC derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing
 CC or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and
 CC anisense therapy. The present sequence is rat hepatoma H35 cell
 CC alpha1-2Fuct catalytic domain.

XX SQ Sequence 353 AA;

Query Match 98.7%; Score 1896; DB 23; Length 353;
 Best Local Similarity 100.0%; Pred. No. 1.3e-198;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 LQORIVKLPSEKELPMTQSSGNTSEPMRRDSEQHNGELRGMTINSIGRLGNQM 65
 DB 1 LQORIVKLPSEKELPMTQSSGNTSEPMRRDSEQHNGELRGMTINSIGRLGNQM 60
 QY 56 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 125
 DB 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
 QY 126 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFRLGLRVNGSOPSTFV 185
 DB 121 YRHIPGHVFRFTGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFRLGLRVNGSOPSTFV 180
 QY 186 GVHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCENINAS 245
 DB 181 GVHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCENINAS 240
 QY 246 RGVVVFAGNIEGSPAKFALLTQCNHIMIGTIGIWAAYLAGGDTIYLYNITLPSDF 305
 DB 241 RGVVVFAGNIEGSPAKFALLTQCNHIMIGTIGIWAAYLAGGDTIYLYNITLPSDF 300
 QY 306 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 358
 DB 301 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF 353

RESULT 5
 AAW37855
 ID AAW37855 standard; Protein; 340 AA.

XX AC AAW37855;

XX DT 28-AUG-1998 (first entry)

XX DE Porcine secretor transferase (FUT2).

XX KW Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;
 transgenic animal; xenotransplantation; organ transplant.

XX OS Sus scrofa.

XX FH Key Location/Qualifiers

FT Domain 1..4

FT Domain /note= "N-terminal cytoplasmic tail"

FT Domain 5..26

FT Domain /note= "transmembrane domain"

FT Domain 27..340

FT Domain /note= "C-terminal domain"

FT Modified-site 185..187

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 251..253

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 279..281

FT Modified-site /note= "Asn is N-glycosylated"

FT Modified-site 305..307

FT Modified-site /note= "Asn is N-glycosylated"

XX WO9807837-A1.

XX PN 26-FEB-1998.

PD

XX 22-AUG-1997; 97WO-AU00540.
 XX PR 23-AUG-1996; 96AU-0001823.
 XX (AUST-) AUSTIN RES INST.
 XX PI McKenzie IFC, Sandrin MS;
 XX WPI; 1998-169148/15.
 DR N-PSDB; AAV29003.
 PT Nucleic acid encoding glycosyltransferase able to compete with
 PT second such enzyme - particularly used to reduce expression of
 PT unwanted carbohydrate epitope(s) on tissues intended for
 PT transplantation
 XX Claim 6; Fig 1A-B; 40pp; English.

XX This polypeptide comprises porcine secretor glycosyltransferase
 CC (SE or FUT2), a type II integral membrane protein has high affinity
 CC for type I and type III substrates. Its amino acid sequence was
 CC deduced from the nucleotide sequence of a genomic DNA clone (see
 CC AAV29003) isolated from a pig liver library on the basis of homology
 CC to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.
 CC Expression of a glycosyltransferase, such as porcine Se, in a
 CC tissue results in reduced expression of unwanted carbohydrate
 CC epitopes on the tissue, especially porcine heart, liver, kidney or
 CC pancreas, rendering it more suitable for transplantation, i.e. less
 CC immunogenic and of increased immunological acceptability. A
 CC claimed method of producing a cell from a donor species that is
 CC immunologically acceptable to a recipient species involves reducing
 CC levels of carbohydrate on the donor cell that causes it to be
 CC recognised as non-self by the recipient by expressing a nucleic
 CC acid for a glycosyltransferase such as porcine Se in the cell.

XX SQ Sequence 340 AA;

Query Match 75.5%; Score 1451; DB 19; Length 340;
 Best Local Similarity 81.9%; Pred. No. 6.8e-150;
 Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHQRIYKVLQPLSEKELPMTTOMSSGNTSEPMRRDSEQHNGELRGMTINSIGR 60
 DB 23 STIHLQORIMVKIQP--TWELQMVQT--TESP-----SSQLKGMWTINA-GR 68
 QY 61 LGNQMGVATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120
 DB 59 LGNQMGVATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 128
 QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFRLGLRVNSQ 180
 DB 129 WMEERYRHIPGEYVRLTGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFRLGLRVNSR 188
 QY 181 PSTFVGHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCRE 240
 DB 189 PSTYGVHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVTSNGMAWCRE 248
 QY 241 NINASRGDVVPAGNIEGSPAKDFALLTQCNTHTIMTIGTIGIWAAYLAGGDTIYLYNITL 300
 DB 249 NINASRGDVVPAGNIEGSPAKDFALLTQCNTHTIMTIGTIGIWAAYLAGGDTIYLYNITL 308
 QY 301 PDSPEFLKFKPEAAFLPEWVGIPADLSPLK 331
 DB 309 PDSPEFLKFKPEAAFLPEWVGIEADLSPLK 339

RESULT 6
 AAW53101.
 ID AAW53101 standard; Protein; 340 AA.

XX AC AAW53101;

XX

DT 08-JUL-1998 (first entry)

XX Pig secretor.

XX Pig; secretor; chimeric; glycosyltransferase; gene therapy;
KW transplantation.

XX Sus scrofa.

XX WO9805768-A1.

XX -12-FEB-1998.

XX 01-AUG-1997; 97WO-AU00492.

XX 21-AUG-1996; 96US-0024279.

XX 02-AUG-1996; 96AU-0001402.

XX (AUST-) AUSTIN RES INST.

XX McKenzie IFC, Sandrin MS;

XX WPI; 1998-159176/14.

XX N-PSDB; AAV21639.

XX Nucleic acids encoding chimeric glycosyltransferases - used for
PT altering carbohydrate levels on the surface of cells, useful in gene
PT therapy and transplantation

PS Example 2; Fig 6; 51pp; English.

XX The present sequence represents pig secretor used in an example of the
CC present invention. The present invention describes nucleic acids (NA)
CC encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
CC domain of a first glycosyltransferase (GT) and a localisation signal of
CC a second GT, whereby when the NA is expressed in a cell and where the
CC chimeric enzyme is located in an area of the cell where it is able to
CC compete for substrate with a second GT, resulting in reduced levels of
CC a product from the second GT. The NAs can be used to produce cells and
CC organs with desired glycosylation patterns. Products and methods of the
CC present invention can be used to reduce the levels of undesirable
CC epitopes in cells, tissues or organs which may be used in
CC transplantation or gene therapy.

XX Sequence 340 AA;

Query Match 75.5%; Score 1451; DB 19; Length 340;

Best Local Similarity 81.9%; Pred. No. 6.8e-150;

Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STTIIHQRIKVLQPISEKELPMTTOMSSGNTSPSPMRDSEOHGNGELRGMTINSIGR 60

DB 23 STIFHLQQRVWKIOP--TWELQWVTQV---TSP-----SSPQLKGMWTINAIGR 68

QY 61 LGNQMGYATLFAARMNGRLAFIPASMINALAPIRISLPVLHSDTAKKIPQWNYHLND 120

DB 69 LGNQMGYATLXALARMNGRPAFIPPEMHSTLAPIRITIPVLHASTARRIPQWNYHLND 128

QY 121 WMEERYRHIPGHEVRTGYPCSWTFYHHLRPELTKETLHDHVRREAQAFLRLGRLYNGSQ 180

DB 129 WMEERYRHIPGHEVRTGYPCSWTFYHHLRPELTKETLHDHVRREAQAFLRLGRLYNGSR 188

QY 181 PSTFVGHVHVRGDIYVHMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240

DB 189 PSTYVGHVHVRGDIYVHMPNVKGVVADRRYLEQAOLDWFRARYRSPVFFVSSNGMAWCRE 248

QY 241 NINASRGDVVFAGNTEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANLYL 300

DB 249 NINASRGDVVFAGNTEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANLYL 308

QY 301 PDSFPLKFKPEAAFLPENWGPADLSPLLK 331

DB 309 PDSFPLKFKPEAAFLPEWIGIEADLSPLLK 339

RESULT 7

AAW69332

ID AAW69332 standard; Protein; 343 AA.

XX AAW69332;

XX 20-NOV-1998 (first entry)

XX Human Sec2 protein sequence.

XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
KW FUT2; nonsecretor genotyping.

XX Homo sapiens.

XX US5807732-A.

XX 15-SEP-1998.

XX 28-FEB-1995; 95US-0395800.

XX 28-FEB-1995; 95US-0395800.

XX (GIOR/) GIORGI D.

XX (KELL/) KELLY R J.

XX (LENN/) LENNON G.

XX (LOWE/) LOWE J B.

XX (ROUQ/) ROUQUIER S.

XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;

XX WPI; 1998-320127/44.

XX N-PSDB; AAV58323.

XX DNA encoding fucosyltransferase enzyme - useful for producing
PT recombinant enzyme and genotyping person as secretor or nonsecretor

XX Disclosure; Column 45-50; 55pp; English.

XX This sequence is the human Sec2 protein of the invention. The DNA
CC encodes a alpha(1,2) fucosyltransferase and is the secretor
CC alpha(1,2)fucosyltransferase locus, that cross hybridises with the
CC H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC individual as a secretor or nonsecretor as it is known that nonsecretors
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC the position corresponding to amino acid 143.

XX Sequence 343 AA;

Query Match 73.3%; Score 1407.5; DB 19; Length 343;

Best Local Similarity 79.2%; Pred. No. 4e-145;

Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STTIIHQRIKVLQPISEKELPMTTOMSSGNTSPSPMRDSEOHGNGELRGMTINSIGR 60

DB 23 STIFHLQQRVWKIOP--TWELQWVTQV---TSP-----SSPQLKGMWTINAIGR 71

QY 61 LGNQMGYATLFAARMNGRLAFIPASMINALAPIRISLPVLHSDTAKKIPQWNYHLND 120

DB 72 LGNQMGYATLXALARMNGRPAFIPPEMHSTLAPIRITIPVLHASTARRIPQWNYHLND 131

QY 121 WMEERYRHIPGHEVRTGYPCSWTFYHHLRPELTKETLHDHVRREAQAFLRLGRLYNGSQ 180

DB 132 WMEERYRHIPGHEVRTGYPCSWTFYHHLRPELTKETLHDHVRREAQAFLRLGRLYNGSR 191

QY 181 PSTFVGHVHVRGDIYVHMPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240

DB 192 PGTFGVGHVVRGDIYVHMPKWKGVVADRRYLQQAOLDWFRARYSSLIFFVVTSSNGMAWCRE 251

Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQMGQYATLALAQLRRAFIIPAMH 117
QY 90 NALAPIFRISLPVLHSDIAKKIPWQNYHLNDWMEERYRHPGHEVFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITLPLVAPEVDSRTPWRQLQHDWSEYADLRDPFLKLSGFPSCSWTFHHL 177
QY 150 REELIKEFTLHDHVRBEAQAFRLGLRV--NGSQPSTFVGVHVRGDDYVHVMPNWKGVYA 207
Db 178 REQIRREFTLHDHREEAQSVLQRLGRTGDRPTFVGVHVRGDDYLVQVMPQKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVYVTSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLRQAMDWFRARHEARVVFVVISNGMEWCKENIDTSQGDVTFAGDGEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGIWAAYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDITVLANFTLPDSEFLKFKPEAAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTLAKP 365

RESULT 10
AAR80154
ID AAR80154 standard; Protein; 365 AA.
XX
AC AAR80154;
XX
DT 25-JUN-1996 (first entry)
XX
DE GDP-L-fucose-beta-D-galactosidase alpha-1,2-fucosyltransferase.
XX
KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase; Fuc-T;
KW synthesis; 2'-fucosyltransferase; oligosaccharide; human milk;
KW non-human transgenic mammal; secondary gene product.
XX
OS Homo sapiens.
XX
PN WO9524494-Al.
XX
PD 14-SEP-1995.
XX
PF 24-JAN-1995; 95WO-US00926.
XX
PR 09-MAR-1994; 94US-0209122.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX
DR WPI; 1995-336739/43.
DR N-PSDB; AAR98461.
XX
XX
PT Prodn. of human-used milk by non-human transgenic mammal - by
PT inserting heterologous gene encoding human catalytic entity which
PT produces oligo:saccharide(s) and glyco:conjugate(s) into mammal's
PT genome
XX
PS Example 1; Page 62-64; 83pp; English.
XX
CC GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase (Fuc-T)
CC is encoded by cDNA (AAQ98461) isolated from a human epidermal carcinoma
CC cell line. The enzyme is responsible for the synthesis of
CC 2'-fucosyllactose, one of the oligosaccharides in human milk. Other
CC products of Fuc-T include glycoproteins contg. beta-linked terminal
CC galactose residues which can be fucosylated by Fuc-T. This DNA can
CC be used to produce non-human transgenic mammals able to produce
CC secondary gene products, e.g. oligosaccharides, in their milk. The
CC transgenic mammals milk biochemically resembles human milk. This
CC humanised milk can be used in the prepn. of an enteral nutritional
CC product useful in the nutritive maintenance of an animal.

XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTESPEMRDSDQHGNGELRGMTFINSICRLGNQMGQYATLALAQLRRAFIIPAMH 89
Db 59 GTAMGNASSCPQH-PASLSGTWTVPNGRFGNQMGQYATLALAQLRRAFIIPAMH 117
QY 90 NALAPIFRISLPVLHSDIAKKIPWQNYHLNDWMEERYRHPGHEVFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITLPLVAPEVDSRTPWRQLQHDWSEYADLRDPFLKLSGFPSCSWTFHHL 177
QY 150 REELIKEFTLHDHVRBEAQAFRLGLRV--NGSQPSTFVGVHVRGDDYVHVMPNWKGVYA 207
Db 178 REQIRREFTLHDHREEAQSVLQRLGRTGDRPTFVGVHVRGDDYLVQVMPQKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVYVTSNGMAWCENINASRGDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLRQAMDWFRARHEARVVFVVISNGMEWCKENIDTSQGDVTFAGDGEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGIWAAYLAGDITIYLANFTLPDSEFLKFKPEAAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDITVLANFTLPDSEFLKFKPEAAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTLAKP 365

RESULT 11
AAR70422
ID AAR70422 standard; Protein; 365 AA.
XX
AC AAR70422;
XX
DT 27-FEB-1996 (first entry)
XX
DE 2-Alpha-fucosyltransferase.
XX
KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
KW glycosyltransferase; glycosylation; oligosaccharide; glycoprotein;
KW glycolipid; transgenic animal; cattle; milk.
XX
OS Homo sapiens.
XX
PN WO9524495-Al.
XX
PD 14-SEP-1995.
XX
PF 24-JAN-1995; 95WO-US00967.
XX
PR 09-MAR-1994; 94US-0208889.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX
DR WPI; 1995-328284/42.
DR N-PSDB; AAT01083.
XX
PT New transgenic non-human mammal milk prods - contg. heterologous
PT components produced as secondary gene prods. of an heterologous gene
XX
PS Example 1; Page 62-64; 83pp; English.
XX
CC 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone
CC (AAR01083) isolated from a human epidermal carcinoma A431 library.
CC The enzyme can be expressed in the milk of a transgenic mammal,
CC esp. cow. This allows large-scale prodn. of oligosaccharides and

CC glycosylated proteins and lipids in the milk.
XX Sequence 365 AA;
SQ Query Match 54.7%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTSEPMRDSQHGNGELRGMTINSICRLGNQMGYATLAFALARMNGRLAFIPASMH 89
DB 59 GTAMGPNASSCPQH-PASISGTTWVYPNGRGNQMGYATLLALAQLNGRAFIIPAMH 117
QY 90 NALAPTRISLPLVLSHDTAKKIPWQNYHLNDWMEERYHIPGHVFTGYPGCSWTFYHHL 149
DB 118 AALAPVFRITLPLVLAPEVDSRTPWRELQHDNMSEYADLRDPFLKLSGFGPCSWTFHHL 177
QY 150 RPEILKEFTLHDHVEEAQAFRLRLV--NGSOPSTFVGVHVRGDDYHVMNPVWKGVA 207
DB 178 REQIRREFTLHDHLEEAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVQMPQKRWGVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMWCARENINASRGDVFVFNAGNIEGSPAKDFALL 267
DB 238 DSAYLRQAMDWRARHEAPVFTVTSNGMEWCKENIDTSQDVTFAAGDGEATPWKDFALL 297
QY 268 TQCNTHITIGTGFVWAAVLGGDTIYLANITLPSDFLKVKPEAAFLPFWGVGIPADLS 327
DB 298 TQCNTHITIGTGFVWAAVLGGDTIYLANITLPSDFLKVKPEAAFLPFWGVGIPADLS 357
QY 328 PLLKALTP 335
DB 358 PLWTLAKP 365
RESULT 12
AAR70421
ID AAR70421 standard; Protein; 365 AA.
AC AAR70421;
XX 27-FEB-1996 (first entry)
DT 2-Alpha-fucosyltransferase.
DE GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
KW transgenic animal; cattle; glycosylation; milk.
XX Homo sapiens.
XX WO9524488-A1.
XX 14-SEP-1995.
PD 24-JAN-1995; 95WO-US01147.
XX 09-MAR-1994; 94US-0209132.
PR (ABBO) ABBOTT LAB.
PA Cummings RD, Koppchik JJ, Moremen KW, Mukerji P;
PI Pierce JM, Prieto PA, Smith DF;
XX WPI; 1995-328279/42.
DR N-PSDB; AAT01082.
XX Transgenic animal expressing heterologous catalyst - used in
PT metabolite prodn esp. glycosyltransferase for prodn. of
PT oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
PT foods, pharmaceuticals, etc.
XX Example 1; Page 62-64; 84pp; English.
PS Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone

CC (AAT01082) isolated from a human epidermal carcinoma cell line A431
XX cDNA library. The enzyme can be expressed in the milk of a
CC transgenic mammal, esp. cow. This allows large-scale prodn. of
CC oligosaccharides or glycosylated proteins and lipids in the milk.
XX Sequence 365 AA;
SQ Query Match 54.7%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTSEPMRDSQHGNGELRGMTINSICRLGNQMGYATLAFALARMNGRLAFIPASMH 89
DB 59 GTAMGPNASSCPQH-PASISGTTWVYPNGRGNQMGYATLLALAQLNGRAFIIPAMH 117
QY 90 NALAPTRISLPLVLSHDTAKKIPWQNYHLNDWMEERYHIPGHVFTGYPGCSWTFYHHL 149
DB 118 AALAPVFRITLPLVLAPEVDSRTPWRELQHDNMSEYADLRDPFLKLSGFGPCSWTFHHL 177
QY 150 RPEILKEFTLHDHVEEAQAFRLRLV--NGSOPSTFVGVHVRGDDYHVMNPVWKGVA 207
DB 178 REQIRREFTLHDHLEEAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVQMPQKRWGVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMWCARENINASRGDVFVFNAGNIEGSPAKDFALL 267
DB 238 DSAYLRQAMDWRARHEAPVFTVTSNGMEWCKENIDTSQDVTFAAGDGEATPWKDFALL 297
QY 268 TQCNTHITIGTGFVWAAVLGGDTIYLANITLPSDFLKVKPEAAFLPFWGVGIPADLS 327
DB 298 TQCNTHITIGTGFVWAAVLGGDTIYLANITLPSDFLKVKPEAAFLPFWGVGIPADLS 357
QY 328 PLLKALTP 335
DB 358 PLWTLAKP 365
RESULT 13
AAR90572
ID AAR90572 standard; Protein; 365 AA.
AC AAR90572;
XX 08-APR-1996 (first entry)
DT Human H-transferase.
DE H-transferase; xenograft hyperacute rejection; transplacental;
KW glycosyltransferase; galactose alpha(1,3) galactose.
XX Homo sapiens.
XX WO9534202-A1.
XX 21-DEC-1995.
PD 14-JUN-1995; 95WO-US07554.
XX 21-JUL-1994; 94US-0278282.
PR 15-JUN-1994; 94US-0260201.
XX (ALEX-) ALEXION PHARM INC.
PA (AUST-) AUSTIN RES INST.
XX Fodor WL, McKenzie IFC, Rother RP, Sandrin MS, Squinto SP;
XX WPI; 1996-049326/05.
DR N-PSDB; AAT12238.
XX Redn. of rejection of xenogeneic cells following transplantation
PT by introducing a vector expressing fucosyltransferase into the
XX cells
XX Example 1; Page 45-47; 69pp; English.

```
XX The human H-transferase (AAK90572) product of a cDNA clone (AAT12236);
CC pred. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
CC can be expressed in xenogeneic organs, tissues and cells using
CC a vector such as PAPX-1 (AAT12239). This results in decreased
CC expression of the non-human antigen galactose alpha(1,3) galactose
CC on the surface of the organs etc. so that hyperacute rejection is
CC reduced upon transplantation to humans.
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 17; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEPMRRDSEQHNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMH 89
Db 59 GTAMGPNASSCCPOH-PASLSGNTWVYPNGRFGNQGYATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYRHPGHVFRFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITPLVLAPEVDSRTPWRELQLHDWMESEYADLRDPFLKLSGFPCCSWTFEHL 177
QY 150 RPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGVHVRGDDYVHVPNWKGVA 207
Db 178 REQIRREFTLHDHLREBEAQSVLGQLRGTRGDRPTFVGHVHVRGDDYLVQVMPQRWKGVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINASRGDDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLQAMDWFRARHEAPVVFVVTSGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKVFKEAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKVFKEAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTLAKP 365
RESULT 14
AAW23805
ID AAW23805 standard; Protein; 365 AA.
AC AAW23805;
XX
XX 15-SEP-1997 (first entry)
XX Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).
XX
XX Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
KW alpha 1,2 FT; transgene; transgenic mouse; animal model;
KW intestinal adhesion; Helicobacter pylori infection; stomach;
KW small intestine; gut; epithelial cell; surface receptor;
KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW gastric adenocarcinoma; Lewis antigen; fucosylation.
XX
OS Homo sapiens.
XX
XX US5625124-A.
XX
XX 29-APR-1997.
XX
XX 11-JUL-1994; 94US-0273411.
XX
XX 11-JUL-1994; 94US-0273411.
XX (UNITW ) UNIV WASHINGTON.
XX
XX Falk P, Gordon JI;
XX
XX WPI: 1997-258275/23.
XX N-PSDB; AAT76768.
```

```
XX Animal model for Helicobacter pylori infection - comprising
PT transgenic mouse expressing human enzyme promoting intestinal
PT adhesion
XX
XX Example 1; Columns 13-16; 24pp; English.
XX
XX A claimed transgenic mouse expresses, in its intestinal epithelial
CC cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
CC fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
CC beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
CC called alpha 1,3/4 FT). The enzyme is expressed under the
CC control of a gut epithelial cell-specific promoter and Helicobacter
CC pylori adheres to the transgenic cells. The transgenic mouse and
CC intestinal epithelial cells from it are useful as models for screening
CC compounds for the ability to inhibit adhesion of H. pylori to gut
CC epithelial cells. The present sequence represents human alpha 1,2 FT
CC and was published by Larsen et al., Proc.Natl.Acad.Sci.USA, 87:
CC 6674-6678 (1990).
XX
SQ Sequence 365 AA;
Query Match 54.7%; Score 1051.5; DB 18; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTPEPMRRDSEQHNGELRGMTFINSIGRLGNOMGEYATLALARMNGRLAFIPASMH 89
Db 59 GTAMGPNASSCCPOH-PASLSGNTWVYPNGRFGNQGYATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYRHPGHVFRFTGYPCSWTFYHHL 149
Db 118 AALAPVFRITPLVLAPEVDSRTPWRELQLHDWMESEYADLRDPFLKLSGFPCCSWTFEHL 177
QY 150 RPEILKEFTLHDHVRREAQAFRLGLRV--NGSOPSTFVGVHVRGDDYVHVPNWKGVA 207
Db 178 REQIRREFTLHDHLREBEAQSVLGQLRGTRGDRPTFVGHVHVRGDDYLVQVMPQRWKGVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINASRGDDVVFAGNGIEGSPAKDFALL 267
Db 238 DSAYLQAMDWFRARHEAPVVFVVTSGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 268 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKVFKEAFLPEWVGIPADLS 327
Db 298 TCNHTIMTIGTFGWAAYLAGDGTIYLANFTLPDSEFLKVFKEAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTLAKP 365
RESULT 15
AAW13640
ID AAW13640 standard; Protein; 365 AA.
XX
XX AAW13640;
XX
XX 19-JUN-1997 (first entry)
XX
XX Human alpha(1,2)-fucosyltransferase.
DE
XX
XX Alpha(1,2)-fucosyltransferase; glycosylation; oligosaccharide;
KW blood group H.
XX
XX Homo sapiens.
XX
XX WO9709421-A1.
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US13616.
XX
XX 08-SEP-1995; 95US-0525058.
XX
```


GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 14.7192 Seconds
(without alignments)
715.625 Million cell updates/sec

Title: US-10-040-863-8_COPY_23_380

Perfect score: 1921

Sequence: 1 STIIHQQRIYVKQLPLSEKE.....RSHFHLKAKGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	360	4	US-09-298-886-8
2	1896	98.7	353	4	US-09-298-886-10
3	1451	75.5	340	4	US-09-254-077A-6
4	1451	75.5	340	4	US-09-254-077A-7
5	1407.5	73.3	343	1	US-08-395-800A-8
6	1407.5	73.3	343	4	US-09-254-077A-8
7	1397	72.7	344	4	US-09-298-886-11
8	1396.5	72.7	347	4	US-09-254-077A-9
9	1086	56.5	373	4	US-09-254-077A-12
10	1051.5	54.7	333	5	PCT-US91-00899-11
11	1051.5	54.7	365	1	US-07-914-281-6
12	1051.5	54.7	365	1	US-08-393-246-6
13	1051.5	54.7	365	1	US-08-273-411-1
14	1051.5	54.7	365	1	US-08-525-058A-6
15	1051.5	54.7	365	1	US-08-395-800A-6
16	1051.5	54.7	365	1	US-08-395-800A-10
17	1051.5	54.7	365	2	US-08-696-731-6
18	1051.5	54.7	365	4	US-09-042-531-6
19	1051.5	54.7	365	4	US-09-254-077A-11
20	1051.5	54.7	365	5	PCT-US91-00899-12
21	1046.5	54.5	365	4	US-09-151-592-2
22	1046.5	54.5	365	4	US-09-254-077A-10
23	696.5	36.3	222	1	US-08-395-800A-2
24	439	22.9	102	1	US-08-395-800A-3
25	141	7.3	300	4	US-09-433-598-2
26	97	5.0	183	1	US-08-167-035-33
27	97	5.0	183	1	US-08-208-887A-33

28.	97	5.0	183	2	US-08-539-005-33	Sequence 33, Appl
29	97	5.0	183	4	US-09-280-598-35	Sequence 35, Appl
30	96.5	5.0	355	4	US-09-390-131-5	Sequence 5, Appl
31	91.5	4.8	801	1	US-07-906-349A-6	Sequence 6, Appl
32	91	4.7	442	1	US-08-220-151-22	Sequence 22, Appl
33	91	4.7	442	1	US-08-413-118-22	Sequence 22, Appl
34	91	4.7	442	3	US-08-473-446-22	Sequence 22, Appl
35	91	4.7	442	6	5470718-3	Patent No. 5470718
36	87	4.5	217	1	US-08-167-035-6	Sequence 6, Appl
37	87	4.5	217	1	US-08-208-887A-6	Sequence 6, Appl
38	87	4.5	217	2	US-08-539-005-6	Sequence 6, Appl
39	87	4.5	217	2	US-08-815-176-3	Sequence 3, Appl
40	87	4.5	217	2	US-08-815-176-4	Sequence 4, Appl
41	87	4.5	217	4	US-08-664-962B-6	Sequence 6, Appl
42	87	4.5	217	4	US-09-311-743-6	Sequence 6, Appl
43	87	4.5	217	4	US-09-280-598-6	Sequence 6, Appl
44	87	4.5	217	4	US-09-197-344-3	Sequence 3, Appl
45	87	4.5	217	4	US-09-197-344-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-298-886-8
; Sequence 8, Application US/09298886
; Patent No. 6329170

GENERAL INFORMATION:

- APPLICANT: Eric H. Holmes et al.
- TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
- TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
- FILE REFERENCE: 8511-029
- CURRENT APPLICATION NUMBER: US/09/298,886
- CURRENT FILING DATE: 1999-04-26
- NUMBER OF SEQ ID NOS: 29
- SOFTWARE: PatentIn Ver. 2.0
- SEQ ID NO 8
- LENGTH: 380
- TYPE: PRT
- ORGANISM: Rattus norvegicus

Query Match 100.0%; Score 1921; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 8.7e-214;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	STIIHQQRIYVKQLPLSEKELPMTQMSNGNTSEPMRRDSEQHNGELRGMTINSIGR	60
DB	23	STIIHQQRIYVKQLPLSEKELPMTQMSNGNTSEPMRRDSEQHNGELRGMTINSIGR	82
QY	61	LGNGMGVATLFAIARNGRLAFTPASHMNAIPIRTSLPVLHSDTAKKIPWONYHLND	120
DB	83	LGNGMGVATLFAIARNGRLAFTPASHMNAIPIRTSLPVLHSDTAKKIPWONYHLND	142
QY	121	WMERYRHPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVFEEAQAFTLRLVRNGSQ	180
DB	143	WMERYRHPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVFEEAQAFTLRLVRNGSQ	202
QY	181	PSTFVGVEVRGDDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMACRE	240
DB	203	PSTFVGVEVRGDDYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMACRE	262
QY	241	NINASRGDVFAGNIGGSPAKDALLTCQNHITMTCTGFIWAAYLAGGDTIYLANVTL	300
DB	263	NINASRGDVFAGNIGGSPAKDALLTCQNHITMTCTGFIWAAYLAGGDTIYLANVTL	322
QY	301	PDSPFLKVFKEPAEAFLEPWGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF	358
DB	323	PDSPFLKVFKEPAEAFLEPWGIPADLSPLKALTPACPRSHFHLKAKGVTCYVAGRAF	380

RESULT 2

US-09-298-886-10
: Sequence 10, Application US/09298886
: Patent No. 6329170
: GENERAL INFORMATION:
: APPLICANT: ERIC H. HOLMES ET AL.
: TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
: TITLE OF INVENTION: GM1-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: 8511-029
: CURRENT APPLICATION NUMBER: US/09/298,886
: CURRENT FILING DATE: 1999-04-26
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Rattus norvegicus
US-09-298-886-10

Query Match 98.7%; Score 1896; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 6.1e-211; Indels 0; Gaps 0;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LQQRIVKQLPLSEKELPMITQMSGNTSEPMRRDSEQHGNGELRGMFTINSIGRLGNQM 65
DB 1 LQQRIVKQLPLSEKELPMITQMSGNTSEPMRRDSEQHGNGELRGMFTINSIGRLGNQM 60

QY 66 GEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEER 125
DB 61 GEYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEER 120

QY 126 YRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQSPSTEV 185
DB 121 YRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQSPSTEV 180

QY 186 GVHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCREINAS 245
DB 181 GVHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCREINAS 240

QY 246 RGVDFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYLTPDSPF 305
DB 241 RGVDFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYLTPDSPF 300

QY 306 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 358
DB 301 LKVPKPEAAFLPEWVGIPADLSPLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3
US-09-254-077A-6
: Sequence 6, Application US/09254077A
: Patent No. 6399758
: GENERAL INFORMATION:
: APPLICANT: SANDRIN, MAURO S.
: APPLICANT: MCKENZIE, IAN C. F.
: TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
: FILE REFERENCE: 30562.5USWO
: CURRENT APPLICATION NUMBER: US/09/254,077A
: CURRENT FILING DATE: 1999-06-11
: PRIOR APPLICATION NUMBER: PCT/AU97/00540
: PRIOR FILING DATE: 1997-08-22
: PRIOR APPLICATION NUMBER: PO 1823
: PRIOR FILING DATE: 1996-08-23
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 340
: TYPE: PRT
: ORGANISM: Sus scrofa
US-09-254-077A-6

Query Match 75.5%; Score 1451; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 2e-159;
Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHLQQRIVKLOPLSEKELPMITQMSGNTSEPMRRDSEQHGNGELRGMFTINSIGR 60
DB 23 STIFHLQQRVWKIOP--TWELQVTVT---TESP-----SSPQLKGMWTINAIGR 68

QY 61 LGNQMGAYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLND 120
DB 69 LGNQMGAYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLND 128

QY 121 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQ 180
DB 129 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSR 188

QY 181 PSTFVGHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 240
DB 189 PSTYGVHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 248

QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYL 300
DB 249 NINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYL 308

QY 301 POSPFLKFKPEAAFLPEWVGIPADLSPLK 331
DB 309 POSPFLKFKPEAAFLPEWVGIPADLSPLK 339

RESULT 4
US-09-254-077A-7
: Sequence 7, Application US/09254077A
: Patent No. 6399758
: GENERAL INFORMATION:
: APPLICANT: SANDRIN, MAURO S.
: APPLICANT: MCKENZIE, IAN C. F.
: TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
: FILE REFERENCE: 30562.5USWO
: CURRENT APPLICATION NUMBER: US/09/254,077A
: CURRENT FILING DATE: 1999-06-11
: PRIOR APPLICATION NUMBER: PCT/AU97/00540
: PRIOR FILING DATE: 1997-08-22
: PRIOR APPLICATION NUMBER: PO 1823
: PRIOR FILING DATE: 1996-08-23
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 7
: LENGTH: 340
: TYPE: PRT
: ORGANISM: Sus scrofa
US-09-254-077A-7

Query Match 75.5%; Score 1451; DB 4; Length 340;
Best Local Similarity 81.9%; Pred. No. 2e-159;
Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHLQQRIVKLOPLSEKELPMITQMSGNTSEPMRRDSEQHGNGELRGMFTINSIGR 60
DB 23 STIFHLQQRVWKIOP--TWELQVTVT---TESP-----SSPQLKGMWTINAIGR 68

QY 61 LGNQMGAYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLND 120
DB 69 LGNQMGAYATLALARMNGRLAFIPASMHNALAPIRISLPVLHSDTAKKIPWQNYHLND 128

QY 121 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSQ 180
DB 129 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGRVNGSR 188

QY 181 PSTFVGHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 240
DB 189 PSTYGVHVVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 248

QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYL 300
DB 249 NINASRGDVFVAGNIEGSPAKDFALLTQCNTHTIMTIGTGWAAAYLAGGDTIYLANTYL 308

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331
Db 309 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 339

RESULT 5
US-08-395-800A-8

; Sequence 8, Application US/08395800A
; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; TITLE OF INVENTION: GENOTYPING A PERSON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLOM, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 343 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-395-800A-8

Query Match

Best Local Similarity 73.3%; Score 1407.5; DB 1; Length 343;

Mismatches 26; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STIIHQORIVKQLPSEKELPMTOMSSGNTSPENRRDSEQHNGELRGMTINSIGR 60

Db 23 STIFVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGR 71

QY 61 LGNOMGEYATLFAALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLND 120

Db 72 LGNOMGEYATLXALAKMNGRPAPFIPQMHSTLAPFRITLPVLHSAFASRIPQNYHLND 131

QY 121 WMEERYRHIPGHVFRGTGYPSCSWTFYHHLRQELQETLHDHVREREAQKFLRGLVNGSQ 180

Db 132 WMEERYRHIPGEYVRETGYPSCSWTFYHHLRQELQETLHDHVREREAQKFLRGLVNGSR 191

QY 181 PSTFVGHVVRGDIYVHVMKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCRE 240

Db 192 PGTFVGHVVRGDIYVHVMKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCRE 251

QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNHMTIMTIGTGIWAAYLAGDITIYLANYYL 300

Db 252 NIDTSHGDVVVAGDIEGSPAKDFALLTQCNHMTIMTIGTGIWAAYLAGDITIYLANYYL 311

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331
Db 312 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 342

RESULT 6

US-09-254-077A-8

; Sequence 8, Application US/09254077A

; Patent No. 6399758

; GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.

; APPLICANT: MCKENZIE, IAN C. F.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES

; FILE REFERENCE: 30562.5USMO

; CURRENT APPLICATION NUMBER: US/09/254,077A

; CURRENT FILING DATE: 1999-06-11

; PRIOR APPLICATION NUMBER: PCT/AU97/00540

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: PO 1823

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-254-077A-8

Query Match

Best Local Similarity 73.3%; Score 1407.5; DB 4; Length 343;

Mismatches 26; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY 1 STIIHQORIVKQLPSEKELPMTOMSSGNTSPENRRDSEQHNGELRGMTINSIGR 60

Db 23 STIFVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTINAIGR 71

QY 61 LGNOMGEYATLFAALRMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLND 120

Db 72 LGNOMGEYATLXALAKMNGRPAPFIPQMHSTLAPFRITLPVLHSAFASRIPQNYHLND 131

QY 121 WMEERYRHIPGHVFRGTGYPSCSWTFYHHLRQELQETLHDHVREREAQKFLRGLVNGSQ 180

Db 132 WMEERYRHIPGEYVRETGYPSCSWTFYHHLRQELQETLHDHVREREAQKFLRGLVNGSR 191

QY 181 PSTFVGHVVRGDIYVHVMKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCRE 240

Db 192 PGTFVGHVVRGDIYVHVMKGVVADRGYLEKALDMFRARYSSPVVVTSGMAWCRE 251

QY 241 NINASRGDVFVAGNIEGSPAKDFALLTQCNHMTIMTIGTGIWAAYLAGDITIYLANYYL 300

Db 252 NIDTSHGDVVVAGDIEGSPAKDFALLTQCNHMTIMTIGTGIWAAYLAGDITIYLANYYL 311

QY 301 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 331

Db 312 PDSPFLKVKPEAAFLPEWVGIPADLSPLLK 342

RESULT 7

US-09-298-886-11

; Sequence 11, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-C29

; CURRENT APPLICATION NUMBER: US/09/298,886

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 344

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-393-246-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTESPMRRDSEQHNGELRGMTTINSIGRLGNOMGEYATLFLALARMGRLAFIPASMH 89
Db 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQGYATLLAQLNGRRRAIFLPAH 117
QY 90 NALAPFRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHL 149
Db 118 AALAPFRITLPVLAPVDSRTPWRELQLEDNWKSEEVADLDFELKLSGPPCSWTFYHHL 177
QY 150 RPELKEFTLHDHVRBAQAFRLGLRV--NGSQPSTFVGHVHVRGDYVHYMPNWKGVVA 207
Db 178 REQIRREFTLHDHVRBAQSVLGQLRLGRGTDRPTFVGHVHVRGDYVHYMPNWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMAMCRNINASRGDVVFNAGNIGSGPAKDFALL 267
Db 238 DSAYLROAMDFRAREAPYFVVTNSGMWCKENIDTSGQDVTFAGDQBATPWKDFALL 297
QY 268 TQCNTHTMTIGTFINAAYLGGDTIYLANITLPDPSFLAVKPEAAFLPEWVGIPADLS 327
Db 298 TQCNTHTMTIGTFINAAYLGGDTVILANFTLPDSEFLKIFKPEAAFLPEWVGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTAKP 365
RESULT 13

US-08-273-411-1
Sequence 1, Application US/08273411
Patent No. 5625124
GENERAL INFORMATION:
APPLICANT: Falk, Per
TITLE OF INVENTION: Animal Model for Gastro-Intestinal
TITLE OF INVENTION: Disease
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,411
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WU106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6508
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..365
OTHER INFORMATION: /note= "GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl
PUBLICATION INFORMATION:
AUTHORS: Larsen, et al.
JOURNAL: Proc. Nat'l Acad. Sci. USA
VOLUME: 87
PAGES: 6674-6678
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365
US-08-273-411-1

Query Match 54.7%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 4.3e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTESPMRRDSEQHNGELRGMTTINSIGRLGNOMGEYATLFLALARMGRLAFIPASMH 89
Db 59 GTAMGNASSCCPQH-PASLSGTWTVPNGRFGNQGYATLLAQLNGRRRAIFLPAH 117
QY 90 NALAPFRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHL 149
Db 118 AALAPFRITLPVLAPVDSRTPWRELQLEDNWKSEEVADLDFELKLSGPPCSWTFYHHL 177
QY 150 RPELKEFTLHDHVRBAQAFRLGLRV--NGSQPSTFVGHVHVRGDYVHYMPNWKGVVA 207
Db 178 REQIRREFTLHDHVRBAQSVLGQLRLGRGTDRPTFVGHVHVRGDYVHYMPNWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMAMCRNINASRGDVVFNAGNIGSGPAKDFALL 267

Db 238 DSAYLRQAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297

QY 266 TQCNHTTMTTGTGWAAYLAGGDTIYLANVTLPDPSFLKVKPEAAFLPEWVGIPADLS 327

Db 298 TQCNHTTMTTGTGWAAYLAGGDTIYLANVTLPDPSFLKVKPEAAFLPEWVGINADLS 357

QY 328 PLKALTP 335

Db 358 PLWTLAKP 365

RESULT 14

US-08-525-058A-6

; Sequence 6, Application US/08525058A

; Patent No. 5770420

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.

; STREET: 1755 Jefferson Davis Highway, Fourth Floor

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,058A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lavalleye, Jean-Paul M. P.

; REGISTRATION NUMBER: 31,451

; REFERENCE/DOCKET NUMBER: 2363-060-55

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)521-4500

; TELEFAX: (703)486-2347

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 365 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-525-058A-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;

Best Local Similarity 63.3%; Pred. No. 4.3e-113;

Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 30 GNTESPWRDSEOHNGELRGMTINSIGRLGNQMGYATLFLALARNGLAFIPASMH 89

Db 59 GTAMGPNASSCCPOH-PASISGTTWTVPNGRFGNQMGYATLLALALQNGRRRAFLPAMH 117

QY 90 NALAPIRISLPVLHSDTAKIPWNYHLNDWMEERYRHIPGHVFRGTGYPSCWTFYHHL 149

Db 118 AALAPVFRITLPVLAPEVDSKTPWRELQLDWNSSEYADLRDPFLKLSGFPSCWTFYHHL 177

QY 150 RPEILKFTLHDHVRERAAQAFRLGLRV--NGSQPSTFGVHVRRGDVYHVMPNWKGVVA 207

Db 178 REQIRREFTLHDHLREERAAQSVLGLRLGRTGDRPTFTGVHVRRGDVYLVQMPQWKGVVG 237

QY 208 DRGYLEKALDMFRARYSSPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 267

Db 238 DSAYLRQAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297

QY 268 TQCNHTTMTTGTGWAAYLAGGDTIYLANVTLPDPSFLKVKPEAAFLPEWVGIPADLS 327

Db 298 TQCNHTTMTTGTGWAAYLAGGDTIYLANVTLPDPSFLKVKPEAAFLPEWVGINADLS 357

QY 328 PLKALTP 335

Db 358 PLWTLAKP 365

RESULT 15

US-08-395-800A-6

; Sequence 6, Application US/08395800A

; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B.

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J.

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 365 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-395-800A-6

Query Match 54.7%; Score 1051.5; DB 1; Length 365;

Best Local Similarity 63.3%; Pred. No. 4.3e-113;

Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 30 GNTESPWRDSEOHNGELRGMTINSIGRLGNQMGYATLFLALARNGLAFIPASMH 89

Db 59 GTAMGPNASSCCPOH-PASISGTTWTVPNGRFGNQMGYATLLALALQNGRRRAFLPAMH 117

QY 90 NALAPIRISLPVLHSDTAKIPWNYHLNDWMEERYRHIPGHVFRGTGYPSCWTFYHHL 149

Db 118 AALAPVFRITLPVLAPEVDSKTPWRELQLDWNSSEYADLRDPFLKLSGFPSCWTFYHHL 177

QY 150 RPEILKFTLHDHVRERAAQAFRLGLRV--NGSQPSTFGVHVRRGDVYHVMPNWKGVVA 207

Db 178 REQIRREFTLHDHLREERAAQSVLGLRLGRTGDRPTFTGVHVRRGDVYLVQMPQWKGVVG 237

QY 208 DRGYLEKALDMFRARYSSPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 267

Db 238 DSAYLRQAMDWFRARHEAPVFTVTSNGMCKENIDTSQGDVTFAGDQGEATPKWKDFALL 297

QY 268 TOCNHTIMTIGTFGIWAAYLAGDITIYLANYTLPDSPELKVPEAEAFLEPWGIPADLS 327
Db 298 TOCNHTIMTIGTFGEFWAAYLAGDITVYLANFTLPDSEELKIKPEAEAFLEPWGINADLS 357
QY 328 PLLKALTP 335
Db 358 PLWTLAKP 365

Search completed: May 27, 2003, 15:11:51
Job time : 15.7192 secs

[illegible]

2058	9	US-09-815-379-17
217	10	US-09-765-298A-6
738	9	US-10-235-521-1
344	9	US-09-895-813A-56
344	10	US-09-815-242-11366
747	9	US-10-028-072-212
747	9	US-10-121-049-212
747	9	US-10-123-904-212
747	9	US-10-140-470-212
747	9	US-10-175-746-212
747	9	US-10-176-918-212
747	9	US-10-176-921-212
747	9	US-10-137-865-212
747	9	US-10-140-474-212
747	9	US-10-142-431-212
747	9	US-10-143-114-212
747	9	US-10-140-002-212
747	9	US-10-142-419-212
747	9	US-10-123-262-212
747	9	US-10-142-423-212
747	9	US-10-121-050-212
747	9	US-10-141-755-212
747	9	US-10-143-032-212
747	9	US-10-123-108-212
747	9	US-10-123-236-212
747	9	US-10-123-261-212

20	89	4.6
21	87	4.5
22	85.5	4.5
23	84.5	4.4
24	84.5	4.4
25	83.5	4.3
26	83.5	4.3
27	83.5	4.3
28	83.5	4.3
29	83.5	4.3
30	83.5	4.3
31	83.5	4.3
32	83.5	4.3
33	83.5	4.3
34	83.5	4.3
35	83.5	4.3
36	83.5	4.3
37	83.5	4.3
38	83.5	4.3
39	83.5	4.3
40	83.5	4.3
41	83.5	4.3
42	83.5	4.3
43	83.5	4.3
44	83.5	4.3
45	83.5	4.3

Q

578
to.

.4676 Seconds
gments)
lion cell updates/

AKRGVTCYVAGRAF 358

375593

```

version 5.1.4_p5
- - 2003 CompuGen

v model

99 ; Search time
      (without a
      2295.367 M
      23_380

XE.....RSHFFH
0.5
5 residues
en parameters:

```

GenCorr
Copyright (c) 1999
in search, using
Y 27, 2003, 15:08
-10-040-863-8_C00
21
STIIHQQRIVKLQPL
OSUM62
pop 10.0 , Gapex
5593 seqs, 991722
ts satisfying ch
length: 0
gth: 200000000

```

OM protein - pro
Run on:
Title:
Perfect score:
Sequence:
Scoring table:
Searched:
Total number of
Minimum DB seq 1
Maximum DB seq 1

```

ALIGNMENTS

RESULT 1
 US-09-999-672-8
 Sequence 8, Application US/09999672
 Patent No. US20020127655A1
 GENERAL INFORMATION:
 APPLICANT: Eric H. Holmes et al.
 TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
 TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 8511-029
 CURRENT APPLICATION NUMBER: US/09/999,672
 CURRENT FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: US/09/298,886
 PRIOR FILING DATE: 1999-04-26
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 8
 LENGTH: 380
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-09-999-672-8

Query Match	100.0%	Score 1921;	DB 10;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 1.7e-190;		
Matches 358;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	STIITHLQORIYKLOPLESEKELPMTQMTQSSGNTSPSPMRDRDSQHGNGELRGMTFINSIGR	60		
23	STIITHLQORIYKLOPLESEKELPMTQMTQSSGNTSPSPMRDRDSQHGNGELRGMTFINSIGR	82		
61	LGNQMGEXATILFALARNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPQWNYHND	120		
83	LGNQMGEXATILFALARNGRLATIPASMHNALAPIFRISLPVLHSDTAKKIPQWNYHND	147		
121	WMEERYRIIPGHFVRFYCYPCSWTFYHHLPEPEILKFTLHDHVREAAQAFRLGRLVNSQ	180		
143	WMEERYRIIPGHFVRFYCYPCSWTFYHHLPEPEILKFTLHDHVREAAQAFRLGRLVNSQ	207		
181	PSTFFGVGHVRRGDVYHVPNPVNWKGVDADGYLEKALDMPFARYSSPFFVVTISNGMAWCRE	240		
203	PSTFFGVGHVRRGDVYHVPNPVNWKGVDADGYLEKALDMPFARYSSPFFVVTISNGMAWCRE	267		

predicted. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	380	10	US-09-999-672-8
2	1921	100.0	380	12	US-10-040-863-8
3	1896	98.7	353	10	US-09-999-672-10
4	1896	98.7	353	12	US-10-040-863-10
5	1451	75.5	340	10	US-09-051-034A-2
6	1397	72.7	344	10	US-09-999-672-11
7	1397	72.7	344	12	US-10-040-863-11
8	1051.5	54.7	365	9	US-10-195-963-10
9	1051.5	54.7	365	10	US-09-863-475A-6
10	1046.5	54.5	365	10	US-09-051-034A-4
11	1044.5	54.4	365	10	US-09-844-258-13
12	1044.5	54.4	365	10	US-09-844-705-13
13	141	7.3	300	10	US-09-848-838-2
14	99	5.2	496	9	US-10-214-524-25
15	97.5	5.1	431	9	US-09-479-614-14
16	97.5	5.1	496	9	US-09-479-614-2
17	97.5	5.1	496	9	US-09-479-614-29
18	89	4.6	1031	9	US-09-764-868-647
19	89	4.6	2057	9	US-09-815-379-10

QY 241 NINASRGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYT 300
Db 263 NINASRGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYT 322
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 358
Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 380

RESULT 2

US-10-040-863-8
Query Match 100.0%; Score 1921; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.7e-190;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Sequence 8, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-8

Query Match 100.0%; Score 1921; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.7e-190;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIIHQRIQVVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGR 60
Db 23 STIIHQRIQVVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGR 82
QY 61 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLND 120
Db 83 LGNOMGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLND 142
QY 121 WMEERYHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 143 WMEERYHIPGHFVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 202
QY 181 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 240
Db 203 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRE 262
QY 241 NINASRGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYT 300
Db 263 NINASRGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYT 322
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 358
Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 380

RESULT 3

US-09-999-672-10
Query Match 98.7%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Sequence 10, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886

; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-10

Query Match 98.7%; Score 1896; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LQORIVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNOM 65
Db 1 LQORIVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNOM 60
QY 66 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLNDWMEER 125
Db 61 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120
QY 126 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSTFV 185
Db 121 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSTFV 180
QY 186 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRENINAS 245
Db 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCRENINAS 240
QY 246 RGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYTLPDSPF 305
Db 241 RGDVVFAGNIEGSPAKDFALLTQCNTIMTIGTGGIWAAYLAGSDTIYLANYTLPDSPF 300
QY 306 LKVKPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 358
Db 301 LKVKPEAAFLPEWVGIPADLSPLKALTACPRSHFLKAKGVTCTYVAGRAF 353

RESULT 4

US-10-040-863-10
Query Match 98.7%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

Query Match 98.7%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

QY 6 LQORIVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNOM 65
Db 1 LQORIVKLOPLSEKELPMTOMSSGNTSEPMRRDSEHQNGELRGMTINSIGRLGNOM 60
QY 66 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLNDWMEER 125
Db 61 GEYATLAFALARMNGRLAFIPASMHNALAPIFRISLPVLSHSDTAKKIPWQNYHLNDWMEER 120
QY 126 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSTFV 185

Db 121 YRHPGHEVFTGTPCWTFFHLLRPEILKSEFTLHDHVREAAQAFRLGLRVNGSQPSTFF 180
QY 186 GVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCRNINAS 245
Db 181 GVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCRNINAS 240
QY 246 RGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANVTLPSDF 305
Db 241 RGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANVTLPSDF 300
QY 306 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHPHLKAQKAGVTCYVAGRAF 358
Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHPHLKAQKAGVTCYVAGRAF 353

RESULT 5
US-09-051-034A-2
; Sequence 2, Application US/09051034A
; Patent No. US2001005584A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL
; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
; FILE REFERENCE: 30562.6USWO
; CURRENT APPLICATION NUMBER: US/09/051,034A
; CURRENT FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: PCT/AU97/00492
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: 60/024,279
; PRIOR FILING DATE: 1996-08-21
; PRIOR APPLICATION NUMBER: P04402
; PRIOR FILING DATE: 1996-08-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus Domesticus
US-09-051-034A-2

Query Match 75.5%; Score 1451; DB 10; Length 340;
Best Local Similarity 81.9%; Pred. No. 6.9e-142;
Matches 271; Conservative 26; Mismatches 20; Indels 14; Gaps 3;

QY 1 STIIHQORIVKQLPSEKELPMTQSSGNTSEPMRDSQHGNGELRCGMFTINSIGR 60
Db 23 STIIFHQORIVKQLPSEKELPMTQSSGNTSEPMRDSQHGNGELRCGMFTINSIGR 68
QY 61 LGNOMGEYATLALARMNGRLAFIPASHHNLALAFIRISLPVLHSDTAKKIPWQNYHLND 120
Db 69 LGNOMGEYATLYALARMNGRPAFIPPMHSTLAFIRITLPLVHASTARRIPWQNYHLND 128
QY 121 WMEERYRHPGHEVFTGTPCWTFFHLLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 129 WMEERYRHPGHEVFTGTPCWTFFHLLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 188
QY 181 PSTVGVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCR 240
Db 189 PSTVGVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCR 248
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANVTL 300
Db 249 NINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANVTL 308
QY 301 PDSFPLKVFKEPAALPEWVGIPADLSPLLK 331
Db 309 PDSFPLKVFKEPAALPEWVGIPADLSPLLK 339

RESULT 6
US-09-999-672-11

; Sequence 11, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-999-672-11

Query Match 72.7%; Score 1397; DB 10; Length 344;
Best Local Similarity 78.9%; Pred. No. 2.8e-136;
Matches 262; Conservative 26; Mismatches 32; Indels 12; Gaps 3;

QY 1 STIIHQORIVKQLPSEKELPMTQSSGNTSEPMRDSQHGNGELRCGMFTINSIGR 60
Db 23 STIIFHQORIVKQLPSEKELPMTQSSGNTSEPMRDSQHGNGELRCGMFTINSIGR 71
QY 61 LGNOMGEYATLALARMNGRLAFIPASHHNLALAFIRISLPVLHSDTAKKIPWQNYHLND 120
Db 72 LGNOMGEYATLYALARMNGRPAFIPPMHSTLAFIRITLPLVHASTARRIPWQNYHLND 131
QY 121 WMEERYRHPGHEVFTGTPCWTFFHLLRPEILKEFTLHDHVREAAQAFRLGLRVNGS 179
Db 132 WMEERYRHPGHEVFTGTPCWTFFHLLRPEILKEFTLHDHVREAAQAFRLGLRVNGS 191
QY 180 QPSTFVGVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCR 239
Db 192 RPTFVGVHVRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFVVTVTSNGMAWCR 251
QY 240 ENINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANV 299
Db 252 ENINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTMTGTFGWAAYLAGDGTIYLYANV 311
QY 300 LPDSFPLKVFKEPAALPEWVGIPADLSPLLK 331
Db 312 LPDSFPLKVFKEPAALPEWVGIPADLSPLLK 343

RESULT 7
US-10-040-863-11
; Sequence 11, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-863-11

Query Match 72.7%; Score 1397; DB 12; Length 344;
Best Local Similarity 78.9%; Pred. No. 2.8e-136;


```
Matches 262; Conservative 26; Mismatches 32; Indels 12; Gaps 3;
QY 1 STIIHQRIYKQLPSKELPMITOMSGNTPESPMRRDSEHQNGELRGMTINSIGR 60
DB 23 STIFVQORLAKIQAM--WELPV-----QIFVLASTSKALGPSQLRGMTINAIGR 71
QY 61 LGNQGEYATLALARMNGREAFIPASMHNAALPIFRISLPLVHSDTAKKIPWQNYHLND 120
DB 72 LGNQGEYATLALAKMNGRPAFIQAOMHSTLAPIFRITLPLVHSTASIRIPWQNYHLND 131
QY 121 WMEERYRHI-PGHVRFYFTGPGCSWTFYHHLRPEILKEFTLHDHVRREAAQFLRGKRVNS 179
DB 132 WMEERYRHIIPGVEYVRFYFTGPGCSWTFYHHLRPEILKEFTLHDHVRREAAQFLRGKRVNS 191
QY 180 QPSTFGVHVVRGDIYVHPNWKGVVADRGYLEKALDMFRARYSPFVVTNSGMWACR 239
DB 192 RPTGTFGVHVVRGDIYVHPNWKGVVADRRYVLRQALDMFRARYSSLFVVTNSGMWACR 251
QY 240 ENINASRGDWFAGNIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANT 299
DB 252 ENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTIGTFGIWAAYLAGGDTIYLYANT 311
QY 300 LPDSPLKVKFPEAAFLPVGWIGIPADLSPLK 331
DB 312 LPDSPLKVKFPEAAFLPVGWIGIPADLSPLK 343
RESULT 8
US-10-105-963-10
; Sequence 10, Application US/10105963
; Publication No. US2003068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Denning, Chris
; APPLICANT: Clark, A. John
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-10
Query Match 54.7%; Score 1051.5; DB 9; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.9e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTESPERRDSEHQNGELRGMTINSIGRGNQGEYATLALARMNGRLAFIPASMH 89
DB 59 GTAMGNASSCCPQH-PASLSGTWTVVYVNGRFGNQGVATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHIIPGHVRFYFTGPGCSWTFYHHL 149
DB 118 AALAPVFRITLPLVLAPEVDSRIPWRELQLDHWMSEYADLRDPFLKLSGFPSCSWTFPHHL 177
QY 150 RPEILKEFTLHDHVRREAAQFLRGRLV--NGSOPSTFVGVHVVRGDIYVHPNWKGVVA 207
DB 178 REQTRREFTLHDHVRREAAQFLRGRLGRTGDRPRTFVGVHVVRGDIYVHPNWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMWACRINASRGDVFAGNIEGSPAKDFALL 267
DB 238 DSAYLRQMDWFRARHAPVYVVTNSGMWCKENIDTSQGDVTFAGDQGTATPKDFALL 297
QY 268 TCNHTTMTIGTFGIWAAYLAGGDTIYLYANTTLPDSPLKVKFPEAAFLPVGWIGIPADLS 327
```

```
DB 298 TCNHTTMTIGTFGIWAAYLAGGDIYLYANFTLPDSEFLKFKPEAAFLPVGWIGINADLS 357
QY 328 PLKALTP 335
DB 358 PLWILAKP 365
RESULT 9
US-09-863-475A-6
; Sequence 6, Application US/09863475A
; Patent No. US20020102688A1
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/863,475A
; APPLICATION NUMBER: 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-863-475A-6
Query Match 54.7%; Score 1051.5; DB 10; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.9e-100;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
QY 30 GNTESPERRDSEHQNGELRGMTINSIGRGNQGEYATLALARMNGRLAFIPASMH 89
DB 59 GTAMGNASSCCPQH-PASLSGTWTVVYVNGRFGNQGVATLLAQLNGRRAFILPAMH 117
QY 90 NALAPIFRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHIIPGHVRFYFTGPGCSWTFYHHL 149
DB 118 AALAPVFRITLPLVLAPEVDSRIPWRELQLDHWMSEYADLRDPFLKLSGFPSCSWTFPHHL 177
QY 150 RPEILKEFTLHDHVRREAAQFLRGRLV--NGSOPSTFVGVHVVRGDIYVHPNWKGVVA 207
DB 178 REQTRREFTLHDHVRREAAQFLRGRLGRTGDRPRTFVGVHVVRGDIYVHPNWKGVVG 237
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSGMWACRINASRGDVFAGNIEGSPAKDFALL 267
```

Db 238 DSAYLRQAMDFAHREAPVFTVTSNGMEWCKENIDTSGDVTYFAGDGOEATPWKDFALL 297
 QY 268 TQCNFTMTTGTGFWAAAYLAGGDIYIANYTLDPSPFLKVPKPAAFPLPWGIPADLS 327
 Db 298 TQCNFTMTTGTGFWAAAYLAGGDIYIANYTLDPSEFLKIFKPAAFPLPWGINADLS 357
 QY 328 PLLKALTP 335
 Db 358 PLWTLAKP 365

RESULT 10

US-09-051-034A-4
 ; Sequence 4, Application US/09051034A
 ; Patent No. US20010055584A1

GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL
 ; TITLE OF INVENTION: SANDRIN, MAURO SERGIO
 ; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
 ; FILE REFERENCE: GLYCOSYLTRANSFERASE
 ; CURRENT APPLICATION NUMBER: 30562.6USWO
 ; CURRENT FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00492
 ; PRIOR FILING DATE: 1997-08-01
 ; PRIOR APPLICATION NUMBER: 60/024,279
 ; PRIOR FILING DATE: 1996-08-21
 ; PRIOR APPLICATION NUMBER: P01402
 ; PRIOR FILING DATE: 1996-08-02
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Sus Domesticus

US-09-051-034A-4

Query Match 54.5%; Score 1046.5; DB 10; Length 365;
 Best Local Similarity 62.6%; Pred. No. 6.2e-100;
 Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPMRDRDSEQHNGELRGMTINSIGRLGNQMGYATLALARMNGRL 81
 Db 51 PVAIFCLAGTPVHPNADSCPKH-PASPSGTWTIYDPGRFGNQMGYATLLALAQLNGRQ 109
 QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141
 Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAPWRELELDHDMSEDAHLKEPWLKLTGFPC 169
 QY 142 SMTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGHVRRGDYVHVMP 199
 Db 170 SMTFFHHLRQIRSEFTLHDHLRQEAQVLSQFLPRTGDRSTFVGHVRRGDYVHVMP 229
 QY 200 NTWKGVADRGYLEKALDMFRARYSSPVFVTSNGMACREINASRGDVVFAGNGIEGS 259
 Db 230 KRWKGVGDGAYLQQAQDMFRARYEAPVFTVTSNGMEWCKENIDTSRGDVIIFAGDGEAA 289
 QY 260 PAKDFALLTQCNTTMTTGTGFWAAAYLAGGDIYIANYTLDPSPFLKVPKPAAFPLPW 319
 Db 290 PARDFALLVQCNTTMTTGTGFWAAAYLAGGDIYIANYTLTSSFLKIFKPAAFPLPW 349
 QY 320 VGIPADLSPL 329
 Db 350 VGINADLSPL 359

RESULT 11

US-09-844-268-13
 ; Sequence 13, Application US/09844268
 ; Patent No. US20020129395A1

GENERAL INFORMATION:

; APPLICANT: BOSWORTH, BRAD

; APPLICANT: VOGELI, PETER
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
 ; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
 ; CURRENT APPLICATION NUMBER: US/09/844,268
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 09/443,766
 ; PRIOR FILING DATE: 1999-11-19
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Porcine
 ; US-09-844-268-13

Query Match 54.4%; Score 1044.5; DB 10; Length 365;
 Best Local Similarity 62.6%; Pred. No. 1e-99;
 Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPMRDRDSEQHNGELRGMTINSIGRLGNQMGYATLALARMNGRL 81
 Db 51 PVAIFCLAGTPVHPNADSCPKH-PASPSGTWTIYDPGRFGNQMGYATLLALAQLNGRQ 109
 QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141
 Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAPWRELELDHDMSEDAHLKEPWLKLTGFPC 169
 QY 142 SMTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGHVRRGDYVHVMP 199
 Db 170 SMTFFHHLRQIRSEFTLHDHLRQEAQVLSQFLPRTGDRSTFVGHVRRGDYVHVMP 229
 QY 200 NTWKGVADRGYLEKALDMFRARYSSPVFVTSNGMACREINASRGDVVFAGNGIEGS 259
 Db 230 KRWKGVGDGAYLQQAQDMFRARYEAPVFTVTSNGMEWCKENIDTSRGDVIIFAGDGEAA 289
 QY 260 PAKDFALLTQCNTTMTTGTGFWAAAYLAGGDIYIANYTLDPSPFLKVPKPAAFPLPW 319
 Db 290 PARDFALLVQCNTTMTTGTGFWAAAYLAGGDIYIANYTLTSSFLKIFKPAAFPLPW 349
 QY 320 VGIPADLSPL 329
 Db 350 VGINADLSPL 359

RESULT 12

US-09-844-705-13
 ; Sequence 13, Application US/09844705
 ; Patent No. US20020133836A1

GENERAL INFORMATION:

; APPLICANT: BOSWORTH, BRAD
 ; APPLICANT: VOGELI, PETER
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
 ; FILE REFERENCE: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
 ; CURRENT APPLICATION NUMBER: US/09/844,705
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 09/443,766
 ; PRIOR FILING DATE: 1999-11-19
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Porcine
 ; US-09-844-705-13

Query Match 54.4%; Score 1044.5; DB 10; Length 365;
 Best Local Similarity 62.6%; Pred. No. 1e-99;
 Matches 194; Conservative 38; Mismatches 75; Indels 3; Gaps 2;

QY 22 PMTQMSSGNTSPMRDRDSEQHNGELRGMTINSIGRLGNQMGYATLALARMNGRL 81
 Db 51 PVAIFCLAGTPVHPNADSCPKH-PASPSGTWTIYDPGRFGNQMGYATLLALAQLNGRQ 109
 QY 82 AFIPASMHNALAPIRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFTGYPC 141
 Db 110 AFIQPAMHVALPVPFRITLPVLAPEVDHRAPWRELELDHDMSEDAHLKEPWLKLTGFPC 169
 QY 142 SMTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRV--NGSQPSTFVGHVRRGDYVHVMP 199
 Db 170 SMTFFHHLRQIRSEFTLHDHLRQEAQVLSQFLPRTGDRSTFVGHVRRGDYVHVMP 229
 QY 200 NTWKGVADRGYLEKALDMFRARYSSPVFVTSNGMACREINASRGDVVFAGNGIEGS 259
 Db 230 KRWKGVGDGAYLQQAQDMFRARYEAPVFTVTSNGMEWCKENIDTSRGDVIIFAGDGEAA 289
 QY 260 PAKDFALLTQCNTTMTTGTGFWAAAYLAGGDIYIANYTLDPSPFLKVPKPAAFPLPW 319
 Db 290 PARDFALLVQCNTTMTTGTGFWAAAYLAGGDIYIANYTLTSSFLKIFKPAAFPLPW 349
 QY 320 VGIPADLSPL 329
 Db 350 VGINADLSPL 359

Db 51 PVAIFCLAGTVPHNDSADSCPKH--PASFSGTWTYPDPGRFGNQMGQYATLALLAQLNGRQ 109
QY 82 APTPASHNALAPIERISLVLHSDTAKKIPWQNYHLNDWMEERYRHPGHFVRFYPC 141
Db 110 AFTOPAMHVALPVRITLVLAPVDRHAPWRELEHDMSEDYAHLKBFWLKLTGTEPC 169
QY 142 SWTFYHHLRPEILKEFTLHDVREAAQAFRLGRV--NGSQPSRFVGVHVRGDDYVHVMP 199
Db 170 SWTFYHHLRQISEFTLHDHRLQEAQGVLSQFLPTGDRPSTFVGVHVRGDDYLRVMP 229
QY 200 NVKGVVADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRENINASRGDVVYFAGNGIEGS 259
Db 230 KRWKGVVGDGRYLLQQAQMDFRARVAFVVFVVTNSNGMEWCRKNIDTSRGDVIFAGDGREAA 289
QY 260 PAKDFALLTQCNTHTIMTIGFGIWAAYLAGDGTIYLANYTLPDPSFLLKVPKPEAAFLPEW 319
Db 290 PARDFALLVQCNHNTIMTIGTGFWAAYLAGDGTIYLANFTLPTSSFLKVPKPEAAFLPEW 349
QY 320 VGIPADLSPL 329
Db 350 VGINADLSPL 359
RESULT 13
US-09-848-838-2
; Sequence 2, Application US/09848838
; Patent No. US20020037570A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcio, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-848-838-2

Query Match 7.3%; Score 141; DB 10; Length 300;
Best Local Similarity 24.0%; Pred. No. 3e-06;
Matches 81; Conservative 33; Mismatches 125; Indels 98; Gaps 14;
QY 59 GRLGNQMGEXA-----TLFALARMNGRLAFIPASMHNALAPIFRIS-- 99
Db 10 GGLGNQMGQFAFAKSLQKHLNTPVLLDTTSDNSNRKAKLELPIDLPYANAKETAKM 69
QY 100 --LPVLHSDTAKKIPWQNYHLNDWMEERYRHI-PGHVFRFTGYPCSWTFYHHLRPEILKE 156
Db 70 QHLFLKLYRDALKYIGEDRVSQEIVFEVEPKLKPRLTYFFGYFQDPDYDAISLLKQT 129
QY 157 FTL-----HDHVREAAQAFRLGRVNGSQPSFTVGVHVRGDDYVHVMPNVKGVVA 207
Db 130 FTLPPLPPPPENKNNKKEEYQKLS--LILAANKSVF--VHIRRGDYVGI-----GCQL 179
QY 208 DRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRENINASRGDVVYFAGNGIEGSPAKDFA-- 265
Db 180 GIDYQKKALEYMAKRVNMEFV-----FCE-----DLKFTQMLDLGYPTDWTTR 225
QY 266 -----LTCQCNHTIMTIGFGIWAAYLAGDGTIYLANYTLPD-----SPFLKVPK 310
Db 226 DKEEYAWDMLMQSCKHGIIANSTYSWAAYL-----MEN--PEKIIIGPKHWF 275
QY 311 PEAFLPEWVGIFADLSPLLKALTTPACPRSHFLKAK 347
+ + + + +
+ + + + +

Db 276 HENLCKRWVKI-----ESHFEVKSQ 296
RESULT 14
US-10-214-524-25
; Sequence 25, Application US/10214524
; Publication No. US20030073142A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Swei-Shen Alex
; APPLICANT: Yang, Yong-Min
; APPLICANT: Barankiewicz, Theresa J.
; APPLICANT: Chen, Zhong
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF
; FILE REFERENCE: ICE-00101.P.1.1
; CURRENT APPLICATION NUMBER: US/10/214,524
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/312,120
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Cat (Felis catus)
US-10-214-524-25

Query Match 5.2%; Score 99; DB 9; Length 496;
Best Local Similarity 25.0%; Pred. NO. 0.14;
Matches 55; Conservative 21; Mismatches 84; Indels 60; Gaps 12;
QY 131 GHFVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRVNGSQPSFTVGVHVR 190
Db 249 GEWVSQKTYTCQYI-----QGTFEDHARKCHESDPRGVSTYLSPPSL----- 293
QY 191 RGOYVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVVTNSNGM--AWCRENINASRGD 248
Db 294 -DLVYHKSPI-TCLVVD-----LANTDGMILTWSREN----- 324
QY 249 VVFNAGNGIEGSPAKDFALLTQCNTHTIMTIGFGIWAAYLAGDGTIYLANYTLPDSP---F 305
Db 325 ----GESVHPDP---MVKKTYNGTITVTSTLPVDATDWVEGET-YOCKVTHPDLPKCI 376
QY 306 LKVEK-PEAAFLPE-WVGIPADLSPLLK-ALTPACPRSHF 342
Db 377 RSIKAPGRPRPPEVYVFLPPEGEPKTKDKVTLTCLQNF 416

RESULT 15
US-09-479-614-14
; Sequence 14, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-14

Query Match 5.1%; Score 97.5; DB 9; Length 431;
Best Local Similarity 25.0%; Pred. No. 0.16;
Matches 52; Conservative 19; Mismatches 78; Indels 59; Gaps 11;

QY 131 GHFVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGRVNGSQPSFTVGVHVR 190

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.9728 Seconds
(without alignments)
2652.940 Million cell updates/sec

Title: US-10-040-863-8_COPY_23_380

Perfect score: 1921

Sequence: 1 STIIHQORIVKLOPLSEKE.....RSHFLKAKGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1407.5	73.3	343	2 A56098	alpha(1,2)fucosylt
2	1296.5	67.5	354	2 B56392	beta-galactoside a
3	1086	56.5	373	2 A56392	beta-galactoside a
4	1051.5	54.7	365	2 A36047	galactoside 2-alph
5	846	44.0	159	2 S46494	galactoside 2-alph
6	581	30.2	142	2 S51582	galactoside 2-alph
7	214.5	11.2	281	2 T44328	hypothetical prote
8	183	9.5	443	2 H87911	protein B0205.4 [i
9	170.5	8.9	363	2 T20745	hypothetical prote
10	154	8.0	348	2 T31916	hypothetical prote
11	153.5	8.0	335	2 T32294	hypothetical prote
12	139.5	7.3	383	2 T20572	hypothetical prote
13	138.5	7.2	500	2 T22068	hypothetical prote
14	133.5	6.9	299	2 H71976	probable alpha(1,2
15	131	6.8	365	2 T23309	hypothetical prote
16	127	6.6	388	2 T32307	hypothetical prote
17	127	6.6	392	2 T23334	hypothetical prote
18	124	6.5	353	2 T25390	hypothetical prote
19	120	6.2	625	2 T21051	hypothetical prote
20	119	6.2	434	2 T26275	hypothetical prote
21	118.5	6.2	381	2 T13140	hypothetical prote
22	117.5	6.1	365	2 T33253	hypothetical prote
23	111	5.8	395	2 T32309	hypothetical prote
24	109.5	5.7	371	2 T32692	hypothetical prote
25	109	5.7	277	2 T25307	hypothetical prote
26	98.5	5.1	317	2 T33887	hypothetical prote
27	96.5	5.0	355	2 T34405	hypothetical prote
28	93.5	4.9	348	2 T02798	hypothetical prote
29	93	4.8	597	2 G87258	acyl-CoA dehydroge

ALIGNMENTS

RESULT 1

A56098

alpha(1,2)fucosyltransferase Sec2, long form - human

C:Species: Homo sapiens (man)

C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 21-Jul-2000

C:Accession: A56098

R:Kelly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.

J. Biol. Chem. 270, 4640-4649, 1995

A:Title: Sequence and expression of a candidate for the human secretor blood group al

tes with the non-secretor phenotype.

A:Reference number: A56098; MUID:95181460; PMID:7876235

A:Accession: A56098

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-343 <REL>

A:Cross-references: GB:U17894; NID:g687618; PIDN:AAC24453.1; PID:g687619

C:Genetics:

A:Gene: GDB:FUT2; SE

A:Cross-references: GDB:120619; OMIM:182100

A:Map position: 19q13.3-19q13.3

C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match 73.3%; Score 1407.5; DB 2; Length 343;
Best Local Similarity 79.2%; Pred. No. 3.3e-110;
Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;

QY	1	STIIHQORIVKLOPLSEKELPMTQSSQNTSEPMRRDSEQHGNGELRGMTINSIGR	60
Db	23	STIFHYQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMNTAIGR	71
QY	61	LGNGMGAYATLPAARMNGRLAIPASMHNALAPIFRISLPVLHSOTAKXIPQWNYHND	120
Db	72	LGNGMGAYATLYALAKHNGRPAFIPAQMHSHTLAPIFRITLPVLHSATASIPQWNYHND	131
QY	121	WMERYRHIPGHEVFRVTGYPCSWTFYHHLRPEILKFTLHDHVREAAQAFRLGRVNGSQ	180
Db	132	WMEEERHIPGHEVFRVTGYPCSWTFYHHLRPEILKFTLHDHVREAAQAFRLGRVNGSR	191
QY	181	PSTFVGHVHRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSSNGMAWCRE	240
Db	192	PGTFVGHVHRRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVFVVTSSNGMAWCRE	251
QY	241	NIWASGDVVVFAGNGEGSPAKDFALLTQCNHMTIMTIGFGIWAAYLAGDGTIYLANVTL	300
Db	252	NIDTSGDVVVFAGNGEGSPAKDFALLTQCNHMTIMTIGFGIWAAYLAGDGTIYLANVTL	311
QY	301	PDSPFLKVKFPEAAFLPEWVGIPADLSPLLK	331
Db	312	PDSPFLKIFKPEAAFLPEWVGIAADLSPLLK	342

RESULT 2

hypothetical prote
transforming prote
probable myp-prote
glycoprotein D pre
fibrillin precurs
probable membrane
peptide synthetase
hypothetical prote
myosin X - human
glycoprotein D pre
modular adaptor Gr
acyl carrier prote
procollagen-lysine
growth factor rece
growth factor rece
NADH2 dehydrogenas

30 92.5 4.8 398 2 T19293
31 92.5 4.8 715 2 T04452
32 92.5 4.8 776 2 E85384
33 91 4.7 442 1 VGBEEA
34 89 4.6 353 2 JN0920
35 89 4.6 468 2 S61964
36 89 4.6 1329 2 AB2934
37 89 4.6 1344 2 E98348
38 89 4.6 2058 2 A59267
39 88.5 4.6 452 1 VGBEG3
40 88 4.6 217 2 A54588
41 88 4.6 350 2 T45082
42 87.5 4.6 365 2 T08679
43 87 4.5 217 2 S26050
44 87 4.5 217 2 A43221
45 87 4.5 438 2 T11903

Biochem. J. 300, 623-626, 1994

7

A;title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially expressed in the developing mouse brain
A;Reference number: S46493; MIMD:S4280382; PMID:8010942
A;Accession: S46494
A;Molecule type: mRNA
A;Residues: 1-159 <PIA>
A;Cross-references: EMBL:L26010; NID:g414816; PIDN:AAB41515.1; PID:g414817
C;Genetics:
A;Gene: FTB
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match	44.0%	Score 846;	DB 2;	Length 159;
Best Local Similarity	99.4%;	Pred. No. 1.3e-63;		
Matches 158;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy 200	NWKGVVADRGYLEKALDMPFARYSSPVFVTSNGMAWCRENINASRGDVVFACNGIEGS	259		
Db 1	NWKGVVADRGYLEKALDMPFARYSSPVFVTSNGMAWCRENINASRGDVVFACNGIEGS	60		
Qy 260	PAKDFALLTCOCHHTIMTIGTGGIWAAYLAGDGTIYLANYTLPDSPFFLVKFKPEAAAFPEW	319		
Db 61	PAKDFALLTCOCHHTIMTIGTGGIWAAYLAGDGTIYLANYTLPDSPFFLVKFKPEAAAFPEW	120		
Qy 320	VGIPADLSPLIKALTACPRSHFHLKAGVTCYVAGRAF	358		
Db 121	VGIPADLSPLIKALTACPRSHFHLKAGVTCYVAGRAF	159		

RESULT 6
S51582
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C;Accession: S51582; S46493
R;Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
submitted to the EMBL Data Library, November 1993
A;Reference number: S51582
A;Accession: S51582
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-142 <PIA>
A;Cross-references: EMBL:L26009; NID:g414814; PIDN:AB41514.1; PID:g554438
R;Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994
A;Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially
expressed in the rat
A;Reference number: S46493; MUID:94280362; PMID:8010942
A;Accession: S46493
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 9-135 <PI2>
A;Cross-references: EMBL:L26009
C;Keywords: glycosyltransferase; hexosyltransferase

[illegible]

RESULT 7
T44328
hypothetical protein wblA [imported] - Vibrio cholerae
C:Species: Vibrio cholerae

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
C;Accession: T44328
C;R;Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda,
Gene 237, 321-332, 1999
A;Title: The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are cl
A;Reference number: 22749; MUID:99453293; PMID:10521656
A;Accession: T44328
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-281 <YAM>
A;Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BAA33632.1; PID:g3721582
A;Experimental source: strain O22
C;Genetics:
A;Note: wblA
C;Superfamily: *Vibrio cholerae* hypothetical protein wblA

Query Match	11.2%; Score 214.5; DB 2;	Length 281;
Best Local Similarity	26.5%; Pred. No. 2.le-10;	
Matches	79; Conservative	50; Mismatches 102; Indels 67; Gaps 14;
<hr/>		
QY	52 MFTTNSICRLGNQGEVATLPALARMG---	RL---APIPASMNALAPITRISLPUHLS 105
	: : :	: : :
Db	1 MVNKISGGLNQLFOYAVGRAIAIQVGPKLDVSAYKNYKLLNG----	YRLDQFNINA 56
<hr/>		
QY	106 DTAKKIPQNXYHL-----	NDMEEERYHI-----PGHFVPEPT 137
	: :	:
<hr/>		
Db	57 DIANED--EIFHLKGSSNRSLRLRGWLKKNTYYAEKORTIYDVSVFMGAPRY---	LD 111
<hr/>		
QY	138 GYPCSWTFYHHLRPEILKEFTLHDHVREAAFLGRLVNGSQSFTEGVHWVRGDYVHV	197
	: : : : : : :	: :
<hr/>		
Db	112 GYWONEQFYSIRAVILQELWPNOPLSINAAA--HQIKI---	QOTHAVSIHVRGDYLN- 165
<hr/>		
QY	198 MPNVKWGVADRGYLEKALDMFRARYSPFPVVVTSGMAWCREMINASRGSDVVFAGNGIE	257
	: : : : : : : :	: : : : :
<hr/>		
Db	166 HPEIG--VLDDIDYKRAVDYIKEIAPVFVFSDVAWCKDNFNFDSVPFIETDTQE	222
<hr/>		
QY	258 GSPAKDPALLTCQHNTIMTCTGTGWAAYL-AGDGTIYLANYT-----	LPSD 303
	: : :	: :
<hr/>		
Db	223 ---IDDMLMCCOCHNIIVANSFSFWAAWLAWSNDVKIVIAKPTMAENPKGYKWWPDS	277

RESULT 8
H87911
protein B0205.4 [imported] - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: H87911
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A/Accession: H87911
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-443 <STO>
A/Cross-references: GB:chr.I; PIDN:AC15988.l; PID:gsl150470; GSPDB:GN00019; CESP:B0203
A/Note: contains weak similarity to fucosyltransferases
C/Genetics:
A/Gene: B0205.4
A/Map position: 1

```

Query Match          9.5%  Score 183;  DB 2;  Length 443;
Best Local Similarity 20.6%  Pred. No. 1.7e-07;
Matches 72;  Conservative 53;  Mismatches 128;  Indels 96;  Gaps 11;

Qy 61  IGNGEYATLFAARMGRLAFIPA--SMENA--LAFIPRLSPVLHSDTAKKI----- 111
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 103  IGNLNFQVAGLSARETGSILLPSTTTLRADFETTFNDSQFVGEDUSRLAEDLN 162
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 112  -----PWQNYHLNDWMEERYRHIPGIVRTGPGCSWTVEVHLRLPILKE-FTLH 160
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```


Db 163 ASKITILTSCAYRNLSLTFPDSRI-----IERIDGYFONFRYFHPDQKIVKLFTEM 216
QY 161 DHVREAOAF-----LRLGR 175
Db 217 DPKRVKSWNTIYNHPTNHRKPEKSTVSGFVTFQLRDFLENVGLSLAVRNAR 276
QY 176 V-----NGSQ-----PSFVGVYVRRGDVHV--MPNVKGVVADRGYLEKAL 216
Db 277 VIETVANDQALEPEEDAFKATMMVGVHHRGMDISMSNRHGHVDVDTIEYKRAI 336
QY 217 DMFRARYSPVFWTNSGMACREINASRGDVFAGNGIEGSPAKDFALLTQCNTHTMT 276
Db 337 QOISKIYENVAETICSDNVAVARRNLKLGKETLHEF--CPGPREVDMAILSCDSVITS 393
QY 277 IGTGFWAAYL--AGGDTIYLANVTLDPSPFLKVKPEAFLPENWGI 322
Db 394 TGTGWSAYLNVNASPDVYYKYHPAPGVMEKMTKTEYFLKSWTAL 442

RESULT 9
T20745
hypothetical protein F11A5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20745
R:Gardner, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: 219319
A:Accession: T20745
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-363 <SWIL>
A:Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5
A:Experimental source: clone F11A5
C:Genetics:
A:Gene: CESP:F11A5.5
A:Map position: 5
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.9%; Score 170.5; DB 2; Length 363;
Best Local Similarity 23.4%; Pred. No. 1.4e-06;
Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;

QY 57 SIGRLGNMGVATLAFALARMNRLA--FIPASMH-----NALAPIRISLPVLHSDT 107
Db 78 SSSRLGNHLFELASVLSISRELQVPTFIENCYHEKMHEDSNLIPGLMNHLLINGSV 137
QY 108 AKKIPQNVH-----LNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 157
Db 138 PSSVRVKFQKCCFTDDPSLLDNYEDEYHLTG-----THYQ-SWKYFSHMRNELI--- 188
QY 158 TLHDHVREAAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVMPNVKGV-VADRGYLEKA 215
Db 189 ---GYLKTENY---MDLPKGGENTFTCVHVRGDGLRV-----GPHVADENPIRS 236
QY 216 LDMFRARYSPVFWTNSGMACREINASRGDVFAGNGIE-----VETGDDYFEMDSLNRPTSKINAFVSQ 278
Db 237 LNLISQVAKRANAT-----VETGDDYFEMDSLNRPTSKINAFVSQ 278
QY 258 GSPAKDFALL--TQCNTHTMTI--GTFGIWAAYLAGDITIYLANVTLDPDSPFLK----- 307
Db 279 NSPADLLYAKSCDVLVLTAAHSTFGWNGVFSKGNRY---YT--DTQFTKDWILETG 333
QY 308 VKPEAARLPEW 319
Db 334 EFISEDYILPHW 345

RESULT 10
T31916
hypothetical protein C17A2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T31916
R:Sammons, L.; Wohldmann, P.;
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C17A2.
A:Reference number: 221098
A:Accession: T31916
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-348 <SAM>
A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00020; CESP:C17A2.4
A:Experimental source: strain Bristol N2; clone C17A2
C:Genetics:
A:Gene: CESP:C17A2.4
A:Map position: 2
A:Introns: 94/3; 133/3; 168/2; 272/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.0%; Score 154; DB 2; Length 348;
Best Local Similarity 22.8%; Pred. No. 3.3e-05;
Matches 63; Conservative 53; Mismatches 102; Indels 58; Gaps 15;

QY 57 STGRGNMGVATLAFALARMNRLA--LAFIPASMHNALAPIRISLP-----VLHSDT 107
Db 36 SPSRLGNHIFEFASLFLSERLHRTPLFLVNEFHOKMLDETRKVMPLGVEKFTVINGSL 95
QY 108 AKKI---PQK-----NYHLNDWMEERYRHIPGHVFTGYPCSWTFYHHLRPEILKEF 157
Db 96 PKSIKQTPFKVCCCRHENPEILEKIDEXYHLTGMYQ-----SWKYFPNMRQELL--- 146
QY 158 TLHDHVREAAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVMPNVKGV-VADRGYLEKAL 216
Db 147 ---DFLDDSSQDF--GNLPRSNQRTHTVTCVHARRGFDV-----GFOAADPDFIRNSV 195
QY 217 DMFRARYSPV-----FVYTSNGMACR---EN--INASRGDVFVAGNGI--EGSPA 261
Db 196 KYIAENFIPEIEYKVKHRRKVVIFGDDLEFMSLFSVSVSTDEPEYMFPAEYIISONSPA 255
QY 262 KDFALLTG--CNHTMTI--GTFGIWAAYLAGDITIY 294
Db 256 EDLTSKQNCDIVLISAPKSTFGWIGYFSKGNKVF 291

RESULT 11
T32294
hypothetical protein K06H6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T32294
R:David, M.; Wohldmann, P.; Bauer, C.; Clarke, K.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid K06H6.
A:Reference number: 221147
A:Accession: T32294
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-335 <DAV>
A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6
A:Experimental source: strain Bristol N2; clone K06H6
C:Genetics:
A:Gene: CESP:K06H6.6
A:Map position: 5
A:Introns: 17/1; 32/2; 58/1; 111/3; 150/3; 184/2; 218/1; 243/3; 272/1; 303/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.0%; Score 153.5; DB 2; Length 335;
Best Local Similarity 22.8%; Pred. No. 3.5e-05;
Matches 77; Conservative 50; Mismatches 118; Indels 93; Gaps 19;

QY 29 SGNIESPENRDSQHGNGELRGMTTINSIGRLGNMGVATLAFALARMNRLA--FIFA 86
Db 27 SATAEIPPLRQESFLIMKRLSRMA--PTARLGNHMFELAAVLGISRLNLTATFTFIED 84

```
QY 87 SMNALAPIERISLP-----VLHSDTAKKIPK--QNYHLN-----DWKEER 125
Db 85 EYRMLTESKEAIPGLVGQFEILNG---KVPYIKNTKLNTRCCVDFVPLIHERNDE 140
QY 126 YRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAEQAFRLGLRVNGSOP---- 181
Db 141 YLHLDGRFYQ-----AWKYPSPMRNELI-----GYLKTSENFGLSPKSNE 180
QY 182 STFYG-VHVRGSDVHVMPNWKGVV-ADRGYLEKALDMF--RARYSSPVFVTSNG---- 234
Db 181 TSFTYCHIRRGDFKRV-----GFAESDEWFIRKAKEFVENKEATSKPCHSVVLFQDDL 234
QY 235 --MAWCREN-----INASRGDVVFAAGNTEGSPAKDFALLTQCNTHTIMTI--GTFGIW 283
Db 235 PFMKNLYENESTHEFVSNSPDDLVIYAKN-----NCDIVLITAPHSSTFGWW 280
QY 284 AAYLAGGDTIYLANV-TLPDSPTLK-VFKPEAAFLPEN 319
Db 281 MGFTSKGDKVYMDIRETRDNVYRNGNLNPDYIYLPWH 318

RESULT 12
T20372
hypothetical protein F08A8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20372
R:Harris, B.
submitted to the EMBL Data Library, October 1997
A:Reference number: Z19294
A:Accession: T20372
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-383 <N1>
A:Cross-references: EMBL:Z99710; PIDN:CAB16868.1; GSPDB:GN00019; CESP:F08A8.5
A:Experimental source: Clone F08A8
C:Genetics:
A:Gene: CESP:F08A8.5
A:Map position: 1
A:Introns: 16/1; 51/3; 77/1; 131/3; 170/3; 209/2; 244/1; 269/1; 297/3; 324/1; 354/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.3%; Score 139.5; DB 2; Length 383;
Best Local Similarity 21.7%; Pred. No. 0.00061;
Matches 81; Conservative 52; Mismatches 135; Indels 105; Gaps 16;

QY 26 QMSGNTSEPMRRDSQHGNGE--LRGMFTINSIGRLGNMGYATLALARNNGRIAF 83
Db 39 QKNSLSPKCEINNSQSPNYEKTQLMLFAFPSPGGIGNKLFEIISLHGATSLQRKAV 98
QY 84 IPA-----SMNALAPIERISLPVLHSD-TAKKIP-WQNYHLNDW----- 121
Db 99 INATNPSEFTLNRIQPLF----PKLADFTLRIIPDSLVTHOOTNWGRCCVYDDPSRF 154
QY 122 --MBERYRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDHVRE-EAQAFRLGLRVNG 178
Db 155 LNRSQDNLILDGHFYQ-----SPKYFHHRPQ-----VREWLAPSKLQAMRAEI 198
QY 179 SQPTF-----GVHVRGSDVHVMPNWKGVV-ADRGYLEKALDMF-RARYSSPVFVTSN 233
Db 199 LLPAKFRDDELICHVVRGSDVHVMPNWKGVV-ADRGYLEKALDMF-RARYSSPVFVTSN 258
QY 234 GMWCRENINASRGDVVFAAGNTEGSPAKDFALLTQCNTHTIMTI--GTFGIW 283
Db 259 -----DIHATVTFEDRVAHFTFLQKPNNSVDYSLPQISPSYTAILPTL 304
QY 270 -----CNHTIMTI--GTFGIWAAYLAGGDT-IYLANV-TLPDSPTLK-VFKPEAA 314
Db 305 TPEIDAFSRLFCOVLTITAPSSTFGWSLVLAKRTATTYRIDLESKDGWAGEMHPEDF 364
QY 315 FLPEWGVIPADLS 327
Db 365 YPPEWIKTKTDLN 377
```

RESULT 13

T22068

hypothetical protein F41D3.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22068

R:White, S.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19508

A:Accession: T22068

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-500 <N1>

A:Cross-references: EMBL:Z81537; PIDN:CAB04377.1; GSPDB:GN00019; CESP:F41D3.6

A:Experimental source: clone F41D3

C:Genetics:

A:Gene: CESP:F41D3.6

A:Map position: 1

A:Introns: 23/1; 59/1; 112/3; 150/3; 222/1; 278/1; 308/3; 326/3; 421/3

Query Match

7.2%; Score 138.5; DB 2; Length 500;

Best Local Similarity 23.4%; Pred. No. 0.001;

Matches 74; Conservative 48; Mismatches 105; Indels 89; Gaps 18;

QY 52 MFTTNSIGRLGNMGYATLALARNNGRIAFIPASMINALAPIR----- 97

Db 49 MTTIVYKGGGQNLFEVLSLGIARKLKRIA-----VFNSSDPVQLQSLNLEFLNKLPRIS 103

QY 98 ---ISLPVLHSDTAK-----KIPWONYHLNDWMEERYRHIPGHVFTGYPCSWTFY 146

Db 104 EQVISVPIEPTETTRFAISSDCCRYEISDNHLD--ESKFLVIEGHFYQ-----SYKIF 155

QY 147 HHLRPEILKEFTLHDHVREAEQAFRLGLRVNGSOPSTFVGHVVRGSDV----HVMPNWW 202

Db 156 ADMKLSI-KEWL----KPEDPEKFRMMISKTESORHK-TCVHVRGSDFLTDEQH----- 203

QY 203 KGVVADRGYLEKALDMF-RARYSSPVFVTSNGMAWCRENINASRGDVVFAAGNTEGSPAK 262

Db 204 --AGTSDSNYITSAIDHLKSLYHGVIFIM-SNDPKWVKVHI-ADHLDY-----QK 248

QY 263 DFALL-----TQCNTHTIMTI--GTFGIWAAYLAGGDT-IYLANV-TLPDS 303

Db 249 DIRIMKLMEDIAIDDLHFSQIYCDVSLITAFSSTFGWSLQSNQSAVYTRDRETKDQ 308

QY 304 PFLKVFKEPAEAFLEPW 319

Db 309 VOLQWTK-EDFYPTW 323

RESULT 14

H71976

probable alpha(1,2)fucosyltransferase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000

C:Accession: H71976

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: H71976

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-299 <ARN>

A:Cross-references: GB:AE001447; GB:AE001439; NID:94154583; PIDN:AAD05659.1; PID:9415

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp0086

C:Superfamily: Vibrio cholerae hypothetical protein wbla

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.478 seconds
(without alignments)
1417.108 Million cell updates/sec

Title: US-10-040-863-8_COPY_23_380

Perfect score: 1921

Sequence: 1 STIIHQRIKVLQPLSEKE.....RSHFHLKAGVTCYVAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query	ID	Description
1	1451	75.5		FUT2_PIG	Q10982 sus scrofa
2	1414	73.6		FUT2_BOVIN	Q28113 bos taurus
3	1407.5	73.3		FUT2_HUMAN	Q10981 h galactosi
4	1338.5	69.7		FUT2_MOUSE	P97353 mus musculu
5	1296.5	67.5		FUT2_RABIT	Q10983 oryctolagus
6	1105.5	57.5		FUT1_MOUSE	Q09160 mus musculu
7	1101.5	57.3		FUT1_RAT	Q10980 rattus norv
8	1086	56.5		FUT1_RABIT	Q10979 oryctolagus
9	1051.5	54.7		FUT1_HUMAN	P19526 homo sapien
10	1050.5	54.7		FUT1_PIG	Q29043 sus scrofa
11	846	44.0		FUT2_RAT	Q10984 rattus norv
12	168	8.7		YK07_CAEEL	P34302 caenorhabdi
13	91	4.7		VGLD_HSVBK	P22484 equine herp
14	89	4.6		FMA3_PORGI	Q51826 porphyromon
15	89	4.6		CDH1_PHACH	Q01738 phanerochaet
16	89	4.6		MY10_HUMAN	Q09467 homo sapien
17	88.5	4.6		PTN7_RAT	P49445 rattus norv
18	88.5	4.6		VGLD_HVEEA	P24872 equine herp
19	88.5	4.6		VGLD_HSVBK	P24379 equine herp
20	88	4.5		GRB2_MOUSE	Q60631 mus musculu
21	87	4.5		GRB2_HUMAN	P29354 homo sapien
22	87	4.5		NOQ1_THETH	Q56222 thermus the
23	87	4.5		QUTD_EWENI	P15325 emericella
24	86.5	4.5		CN3B_HUMAN	Q13370 homo sapien
25	85.5	4.5		PL03_HUMAN	Q60568 homo sapien
26	85.5	4.5		LPH_RAI	Q02401 rattus norv
27	84.5	4.4		RIBB_HELPY	C25484 helicobacte
28	84.5	4.4		AAT_THBAQ	Q33822 thermus aqu
29	84.5	4.4		CPTC_MYCTU	P96875 mycobacteri
30	84	4.4		POIG_MVEV	P03769 m genome po
31	83.5	4.3		TCIA_CAEEL	P03934 caenorhabdi
32	83.5	4.3		FMB2_PORGI	Q51825 porphyromon
33	83.5	4.3		CYB_PARDE	P05418 paracoccus

RESULT 1

ID	FUT2_PIG	STANDARD;	PRT;	340 AA.
AC	Q10982: Q29044; Q19100;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)			
DE	(Fucosyltransferase 2).			
GN	FUT2.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97468270; PubMed=9321466;			
RA	Meijerink E., Fries R., Voegeli P., Masabanda J., Wigger G.,			
RA	Stricker C., Neuenschwander S., Bertschinger H.U., Stranzinger G.;			
RT	"Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6g11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECP18R) loci.";			
RT	Mamm. Genome 8:736-741(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	Cohnsey S., Moutouris E., McKenzie I.F.C., Sandrin M.S.;			
RT	"Molecular cloning and characterization of the pig secretor type alpha(1,2)fucosyltransferase.";			
RT	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
RL	[3]			
RN	SEQUENCE OF 69-334 FROM N.A.			
RP	Petit J.M.;			
RA	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.			
RL	[4]			
RP	SEQUENCE OF 69-80; 119-133 AND 316-334.			
RC	TISSUE=Submaxillary gland;			
RX	MEDLINE=96064742; PubMed=7592879;			
RA	Thurin J., Blaszczyk-Thurin M.;			
RT	"Porcine submaxillary gland GDP-L-fucose: beta-D-galactoside alpha-2-L-fucosyltransferase is likely a counterpart of the human Secretor gene-encoded blood group transferase.";			
RL	J. Biol. Chem. 270:26577-26580(1995).			
CC	-!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R +			
CC	alpha-L-fucosyl-1,2-beta-D-galactosyl-R.			
CC	-!- PATHWAY: Glycosylation.			
CC	-!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			

Q61409 mus musculu
P76484 escherichia
P45784 vibrio chol
P72933 synecocyst
P87503 human adeno
P97846 rattus norv
P78357 homo sapien
P83739 xenopus lae
Q92140 helicobacte
P23739 rattus norv
P47527 mycoplasma
O54991 mus musculu

1 CN3B_MOUSE
1 YFAS_ECOLI
1 GSPN_VIBCH
1 GLO2_SYNY3
1 DPOI_ADE04
1 CTAL_RAT
1 CTAL_HUMAN
1 GRB2_XENLA
1 RIBE_HELPJ
1 SUIS_RAT
1 Y285_MYCGE
1 CTAL_MOUSE

799
1534
252
257
1193
1381
1384
217
344
1840
347
1385

ALIGNMENTS

RT "Sequence and expression of a candidate for the human Secretor blood
RT group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an
RT enzyme-inactivating nonsense mutation commonly correlates with the
RT non-secretor phenotype.";
RL J. Biol. Chem. 270:4640-4649(1995).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT PHE-140.
RX MEDLINE=96199252; PubMed=8621666;
RA Kudo T., Iwasaki H., Nishihara S., Shinya N., Ando T., Narimatsu I.,
RA Narimatsu H.;
RT "Molecular genetic analysis of the human Lewis histo-blood group
RT system. II. Secretor gene inactivation by a novel single missense
RT mutation A385T in Japanese nonsecretor individuals.";
RL J. Biol. Chem. 271:9830-9837(1996).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT PHE-140.
RX MEDLINE=97363210; PubMed=9219535;
RA Koda Y., Soejima M., Wang B., Kimura H.;
RT "Structure and expression of the gene encoding secretor-type
RT galactoside 2-alpha-L-fucosyltransferase (FUT2).";
RL Eur. J. Biochem. 246:750-755(1997).
RN [4]
RP VARIANTS VAL-25; CYS-138 AND ASN-172.
RX MEDLINE=98431007; PubMed=9760207;
RA Liu Y., Koda Y., Soejima M., Pang H., Schlaphoff T., du Toit E.D.,
RA Kimura H.;
RT "Extensive polymorphism of the FUT2 gene in an African (Xhosa)
RT population of South Africa.";
RL Hum. Genet. 103:204-210(1998).
CC -1- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMS FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY THE USE OF
CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC -1- TISSUE SPECIFICITY: SMALL INTESTINE, COLON AND LONG.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U17894; AAC24453.1; -
CC EMBL; D89327; BAA13944.1; -
CC EMBL; D87942; BAA21684.1; -
CC MIM; 182100; -
CC InterPro; IPR002516; GT_11.
CC Pfam; PF01531; Glyco_transf_11; 1.
CC Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Goldi stack; Alternative initiation;
CC Blood group antigen; Polymorphism.
FT CHAIN 1 343 GALACTOSIDE 2-L-FUCOSYLTRANSFERASE 2,
FT CHAIN LONG ISOFORM.
FT CHAIN 12 343 GALACTOSIDE 2-L-FUCOSYLTRANSFERASE 2,
FT CHAIN SHORT ISOFORM.
FT INIT_MET 12 12 FOR SHORT ISOFORM.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 29 343 LUMENAL, CATALYTIC (POTENTIAL).

FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 25 25 I -> V (IN XHOSA POPULATION).
FT VARIANT 138 138 /FTID=VAR_003422.
FT VARIANT 140 140 R -> C (IN XHOSA POPULATION).
FT VARIANT 172 172 /FTID=VAR_003423.
FT VARIANT 343 AA; 39017 MW; 120659CF175E13A CRC64;
SQ SEQUENCE 343 AA; 39017 MW; 120659CF175E13A CRC64;
Query Match 73.3%; Score 1407.5; DB 1; Length 343;
Best Local Similarity 79.2%; Pred. No. 9e-112;
Matches 262; Conservative 26; Mismatches 32; Indels 11; Gaps 2;
QY 1 STIHLQQRIVKQLPLSEKELPMTTSMSSGNSTESPEMRDSEQHNGELGRGFTINSIGR 60
DB 23 STIPHVQRLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTWINAIGR 71
QY 61 LGNOMGEYATLFPALARNNGRLAFIPASMHNNALAFIPRISLPVLESOTAKKIPWONYHND 120
DB 72 LGNOMGEYATLYALAKMNGRPAFIPAQMHSTLAFIPRITLPLVLSATASRIPEWONYHND 131
QY 121 WMEERYRHIFGHFVFTGYPGCSWTFYHHLRPEILKEFTLHDHVREEAQFLRGVRVNGSQ 180
DB 132 WMEERYRHIFGHFVFTGYPGCSWTFYHHLRPEILKEFTLHDHVREEAQFLRGVRVNGSR 191
QY 181 PSTFVGHVHVRGDIYHVMNPVWKGVDGRVLEKALDMFRARYSPFVFTVTSNGMAMCRE 240
DB 192 PGTEVGVHVRGDIYHVMNPVWKGVDGRVLEKALDMFRARYSPFVFTVTSNGMAMCRE 251
QY 241 NINASRGDVFAGNGIEGSPAKDFALLTQCNTTMTTCTFOIWAYLAGGTYILANYTL 300
DB 252 NIDTSHGDFVAGDIEGSPAKDFALLTQCNTTMTTCTFOIWAYLAGGTYILANYTL 311
QY 301 PDSPLKVKPEAAFLPEWVGIPADLSPLK 331
DB 312 PDSPLKIFKPEAAFLPEWVGIAADLSPLK 342
RESULT 4
FUT2_MOUSE
ID FUT2_MOUSE STANDARD; PRT; 368 AA.
AC P97353;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-D-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)
DE {Fucosyltransferase 2} (FUT-III).
GN FUT2 OR SECI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of a mouse GDP-L-Fucose: beta-D-
RT galactoside 2-alpha-L-fucosyltransferase.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Gastrointestinal tract;
RA Lin B., Hayashi Y., Saito M., Sakakihara Y., Yanagisawa M.,
RA Iwamori M.;
RT "Molecular cloning and expression of a GDP-L-fucose:beta-D-galactoside
RT 2-alpha-L-fucosyltransferase in murine gastrointestinal tract.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

```

CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SECL1) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y09882; CAA71008.1; -.
DR EMBL: AF113532; AAD25351.1; -.
DR MGI: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco.transf.11; 1.
DR Transferase: Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
KW DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 42 368 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 368 AA; 41454 MW; 4093E853EB37303B CRC64;

Query Match 69.7%; Score 1338.5; DB 1; Length 368;
Best Local Similarity 77.3%; Pred. No. 6.7e-106;
Matches 252; Conservative 21; Mismatches 36; Indels 17; Gaps 2;

QY 1 STIHLQORIVKLOPLSEKELPMTQMSSGNTSPERRDSEQHNGELRMGTINSIGR 60
||| | : : : : : : : : : : : : : : : : : : : : : : : :
D 36 STIFCHRRLL-----GLVPAPWASPSLVVFPRRHPRE--GMTTIRVKGR 78

QY 61 LGNOMGEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPQNYHLND 120
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 79 LGNOMGEYATLALARMNGRLAFIPASMHSTLAFIPRISLPVLHSDTAKKIPQNYHLND 138

QY 121 WMEERYRHIPGHVFRFGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFGLRUVNGSQ 180
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 139 WMEERYRHIPGHVFRFGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFGLRUVNGSQ 198

QY 181 PSTFVGHVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCR 240
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 199 PSTFVGHVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCR 258

QY 241 NINASRGDVFVAGNGIEGSPAKDFALLTQCNIHTMTGTGFINAAYLAGGDTIYLAN 300
||| : : : : : : : : : : : : : : : : : : : : : : : :
D 259 SITASRGDVFVAGNGIEGSPAKDFALLTQCNIHTMTGTGFINAAYLAGGDTIYLAN 318

QY 301 PDSFELKVFKEAFLPEWVGIPADL 326
| : : : : : : : : : : : : : : : : : : : : : : : :
D 319 PDSFELKVFKEAFLPEWVGIPADL 344

RESULT 5
FUT2_RABIT
ID FUT2_RABIT STANDARD; PRT; 354 AA.
AC Q10983;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (pC 2.4.1.69) (Secretor blood
DE group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-
DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)
DE (Fucosyltransferase 2).

```

```

GN FUT2 OR SECL1 OR RET-II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
RT galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80225; CAA56512.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco.transf.11; 1.
DR Transferase: Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
KW DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 44 354 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6CCB4 CRC64;

Query Match 67.5%; Score 1296.5; DB 1; Length 354;
Best Local Similarity 73.9%; Pred. No. 2.3e-102;
Matches 244; Conservative 25; Mismatches 44; Indels 17; Gaps 2;

QY 1 STIHLQORIVKLOPLSEKELPMTQMSSGNTSPERRDSEQHNGELRMGTINSIGR 60
||| | : : : : : : : : : : : : : : : : : : : : : : : :
D 38 STIFCHRRLL-----VPAWPASRVVVP-----GHLPRGWTINAMGR 80

QY 61 LGNOMGEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPQNYHLND 120
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 81 LGNOMGEYATLALARMNGRLAFIPASMHSTLAFIPRISLPVLHSDTAKKIPQNYHLND 140

QY 121 WMEERYRHIPGHVFRFGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFGLRUVNGSQ 180
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 141 WMEERYRHIPGHVFRFGYPCSWTFYHHLRPEILKFTLHDHVREEAQAFGLRUVNGSR 200

QY 181 PSTFVGHVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCR 240
||||| : : : : : : : : : : : : : : : : : : : : : : : :
D 201 PSTFVGHVRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVVTSGNMAWCR 260

QY 241 NINASRGDVFVAGNGIEGSPAKDFALLTQCNIHTMTGTGFINAAYLAGGDTIYLAN 300
||| : : : : : : : : : : : : : : : : : : : : : : : :

```



```
Db 261 NIDASRGDVVFAGNGLESPAKDFALLTQCNHTYKT~GTGFWAAYLTGGDTVYLANITA 320
QY 301 PDSPFLVFKPEAFIPENTWGPADLSPLL 330
Db 321 PDSPFLVFKPEAFIPENTWGPADLSPLL 350
RESULT 6
FUTL_MOUSE
ID FUTL_MOUSE STANDARD; PRT; 376 AA.
AC 009160;
DT 15-DEC-1998 (Rel. 37, Created);
DT 15-DEC-1998 (Rel. 37, Last sequence update);
DT 15-JUN-2002 (Rel. 41, Last annotation update);
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)F1)
DE (Fucosyltransferase 1).
GN FUTL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss;
RX MEDLINE=97454449; PubMed=9355741;
RA Domino S.E., Hiraiwa N., Lowe J.B.;
RT "Molecular cloning, chromosomal assignment and tissue-specific
RT expression of a murine alpha(1,2)fucosyltransferase expressed in
RT thymic and epididymal epithelial cells.";
RL Biochem. J. 327:105-115(1997).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: IN THE ADULT, HIGHLY EXPRESSED IN PANCREAS,
CC TESTIS AND EPIDIDYMS AND TO A LESSER EXTENT IN THYMUS, LUNG,
CC STOMACH, SMALL INTESTINE, COLON, SPLEEN AND UTERUS. NOT EXPRESSED
CC IN BRAIN, HEART, SKELETAL MUSCLE, KIDNEY, LIVER AND BONE MARROW.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U90553; AAC53492.1; -
DR MGB; MGI:109375; Futl.
DR InterPro; IPR002516; GP_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1
FT TRANSMEM 9 26 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 376 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT FT CARBOHYD 67 67 LUMENAL, CATALYTIC (POTENTIAL).
FT FT CARBOHYD 302 302 N-LINKED (GLCNAC... ) (POTENTIAL).
FT FT CARBOHYD 328 328 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 376 AA; 42255 MW; 21FID24CFE2C4106 CRC64;
Query Match 57.5%; Score 1105.5; DB 1; Length 376;
Best Local Similarity 52.8%; Pred. No. 3.6e-85;
Matches 208; Conservative 38; Mismatches 76; Indels 9; Gaps 2;
QY 15 PLSEKEPLPMTOMSSGNTSEPMRDSQHGELRGMTINSIGLGNOMGEYATLAL 74
Db 53 FVAMVLPFLPLOTSGNSGPSCEQ-----SSLSGTVITTPGFRGNQMGQATLAL 104
```

```
QY 75 ARMGRLAFIPASMHNALAPIERISLVHESDTAKKIPWQNYHLNDWMEERYRHPGHV 134
Db 105 AOLNGRQAFIQEMHAALAPVFRISLPYLDPEVDSLTPWOHLVLDHNMSEYSHLEDPFL 164
QY 135 RFTGPGSWTFYHHLRPELKEFTLHDHVEEAQAFRLGRVNGSQPS-TFVGYHVVRGD 193
Db 165 KLSGFPSCSWTFHHLRPELKEFTLHDHVEEAQAFRLGRVNGSQPS-TFVGYHVVRGD 224
QY 194 YVHVPNVWGVVADRGYLEKALDMFRARYSSPVFVVISNGMAWCRNINASRGDVFVAG 253
Db 225 YLEVMPNRWGVGDRAVYLOQAMDWRARHKDPFVVTSGMKWCLNIDTSHGDDVFVAG 284
QY 254 NGIEGSPAKDFALLTQCNHTIMTIGTFCIWAAYLAGDITVYLANITLPDSPFLVKVPEA 313
Db 285 NGEGTIPGKDFALLTQCNHTIMTIGTFCIWAAYLAGDITVYLANITLPDSEFLKIFRPEA 344
QY 314 AFLPEWVGIPADLSPLLKALTACPRSRPHL 344
Db 345 AFLPEWVGINADLSPLQAQDFPKPKPSLFL 375
RESULT 7
FUTL_RAT
ID FUTL_RAT STANDARD; PRT; 376 AA.
AC Q10980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)F1)
DE (Fucosyltransferase 1).
GN FUT1 OR FTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE-Colon cancer;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-
RT fucosyltransferase genes.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE OF 169-310 FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=94280382; PubMed=8010942;
RA Piau J.-P., Laharriere N., Dabouis G., Denis M.G.;
RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
RT differentially expressed throughout the rat colon.";
RL Biochem. J. 300:623-626(1994).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015637; BAA31130.1; -
DR EMBL; AB006137; BAA21741.1; -
```


DR EMBL; L26009; AAB41514.1; -.
 DR InterPro: IPR002516; GT 11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack.
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 30 376 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 376 AA; 42416 MW; C350C737C758B7F8 CRC64;
 Query Match 57.3%; Score 1101.5; DB 1; Length 376;
 Best Local Similarity 62.3%; Pred. No. 7.8e-86;
 Matches 207; Conservative 37; Mismatches 77; Indels 11; Gaps 2;
 QY 15 PLSEKELPMTOMSSNGTESPEMRDSEHQNGELRGMTINSIGRLGNOMGEYATLAL 74
 Db 53 PVAMVCLPYSNASSGSPSCPE-----QSLSGTWTITPGRGNGMGQYATLAL 103
 QY 75 ARNGRLAFTPASWNAHALAFIRISLPVLSHDSRTAKIPWNYHLNDWMEERYRHIPGHEV 134
 Db 104 AOLNGRAFTQPEMHTTAPVFRISLPVLDPEVDSLTPWQHLVLDWMSSEYSHLEDPEL 163
 QY 135 RFTGPCSWTFYHHLRPEILKEFTLHDVREFAQAFLGLRVN--GSQPSFTFVGVHVRG 192
 Db 164 KLSGFCFSWTFHHLREQIRREFTLHDHLDREDAQRLSLGLRGIPAGIRPIYVGVHVRG 223
 QY 193 DYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVVT-SNGMAWCKRENINASRGDVVFA 252
 Db 224 DYLEVMPNRKGVVGDVAYLQKAMDWFRARHKDPIFVVT-SNGMWGLENDTSHGDVYFA 283
 QY 253 GNGIESPAKDFALLTQCNTHTMTIGTFINWYLAGGDTIYLYANTLPDSEPLKVFKEPE 312
 Db 284 GNGOETPGKDFALLTQCNTHTMTIGTFWYLAGGDTIYLYANTLPDSEPLKVFKEPE 343
 QY 313 AAFPLPWVGIPADLSPLKALTPACPRSHFL 344
 Db 344 AAFPLPWVGINADLSPLQAFDQFDEWDSFLRL 375
 RESULT 8
 FUT1_RABIT
 ID FUT1_RABIT STANDARD; PRT; 373 AA.
 AC Q10979;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
 DE D-galactoside 2-alpha-L-fucosyltransferase 1 (Alpha(1,2)FT 1)
 DE (Fucosyltransferase 1).
 GN FUT1 OR RFT-I.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95238380; PubMed=7721792;
 RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
 RT "Molecular cloning and expression of two types of rabbit beta-galactoside alpha 1,2-fucosyltransferase.";
 RL J. Biol. Chem. 270:8844-8850(1995).
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-1-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-1-R.
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE

CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
 CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
 CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X80226; CAA56513.1; -.
 DR InterPro: IPR002516; GT 11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack.
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 30 373 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 373 AA; 42098 MW; 0A47A1786231525C CRC64;
 Query Match 56.5%; Score 1086; DB 1; Length 373;
 Best Local Similarity 58.7%; Pred. No. 1.6e-84;
 Matches 206; Conservative 50; Mismatches 59; Indels 36; Gaps 5;
 QY 1 STTIHQQRIVK-----LOPLSEKE-----LPMITQMS-----SGNTSPEMRD 40
 Db 23 SFLLHLHQDLFRNGLALSPLCLERQVPAPVAIVCLPVTSPASNASSCAGRPAAPS--- 78
 QY 41 SEQHNGSELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMNALAFIRISL 100
 Db 79 -----GIWTHPDGRFGNQMGQYATLLALQNGRRARFILPAMHAALAPVERITL 128
 QY 101 PVLHSDTAKKIPWNYHLNDWMEERYHPIGVHVRGTVGTPCSWTFYHHLRPEILKEFTLH 160
 Db 129 PVLAEVNRRTSWKQLLHDHWSSEYSLRLEDPFKTGFPCSWTFHHLRPEIRREFTLH 188
 QY 161 DHVREAAQFLRGLRV--NGSQPSIFGVHVRGDIYVHVPNVKGVVADRGYLEKALDM 218
 Db 189 DHLREEAQLLKLRLGRTGARPTFVGVEVRGDIYLVMPQKRGVGDVAYLQAMDW 248
 QY 219 FRARYSSPVFVVTSGMAWCKRENINASRGDVYFAGNGIEGSPAKDFALLTQCNTHTIG 278
 Db 249 FRARHAPIFVVTSGMKWCWENIDASRGDVYFAGNGLESPPAKDFALLTQCNTHTIG 308
 QY 279 TFGIWAAYLAGGDTIYLYANTLPDSEPLKVFKEPEAAFLPEWVGIPADLSPL 329
 Db 309 TEGFWAAYLAGGDTIYLYANTLPDSEPLKVFKEPEAAFLPEWVGINADLSPV 359
 RESULT 9
 FUT1_HUMAN
 ID FUT1_HUMAN STANDARD; PRT; 365 AA.
 AC P19526;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
 DE D-galactoside 2-alpha-L-fucosyltransferase 1 (Alpha(1,2)FT 1)
 DE (Fucosyltransferase 1) (Blood group H alpha 2-fucosyltransferase).
 GN FUT1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.


```
YK07_CAEEL
ID YK07_CAEEL STANDARD; PRT; 365 AA.
AC P34302;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative glycosyl transferase C06E1.7 in chromosome III (EC 2.-.-.-).
GN C06E1.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA LaReille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RN Nature 368:32-38(1994).
[2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L16559; AAA27932.2; -
CC WormPep; C06E1.7; CE30483.
CC InterPro; IPR002516; GT_11.
CC Pfam; PF01531; Glyco.transf_11; 1.
CC Hypothetical protein; Transferase; Glycosyltransferase.
CC KQ SEQUENCE 365 AA; 41991 MW; B5FCA363F31977F CRC64;
CC
CC Query Match 8.7%; Score 168; DB 1; Length 365;
CC Best Local Similarity 22.0%; Pred. No. 7.2e-07;
CC Matches 71; Conservative 48; Mismatches 115; Indels 88; Gaps 16;
CC
CC YQ 57 SIGRLGNMGEXATLALARMGR--LATIPASMHNALAPIFRISLPVLHSD----- 106
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 70 STARLANHIFELVYGMKSLNRKPAIFVEDSKVLLITGVKVLPLGLDEQIFEPY 129
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 107 --TAKKIP-----WQNYHLNDWMEERYRHIPGHEVRTGYPCSWTFYHLLRPILKEF 157
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 130 HNKATKVPILSEKCCIFDNPDKFNISSEVILHLGHFYQ-----SWKYF----- 172
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 158 TLHDHVRBAQAFRLGL-----RVNGSQPSTFVG---VHVRGDIY---HVMPNWKGVA 207
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 173 ---DKYKEKQSVKPAIDFSLPNSDSSNFISRICIHRIRDFDVGQHSNV----- 223
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 208 DRGYLEKALDMFAR-----YSSPVFVTSNGMWCNEMINASRGD 248
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC Db 224 --SFIKPALEFKEREQKDVNKKMLTVIMGDDPDFEAKMF-----EGTVRAKKEAKEETT 277
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
CC YQ 249 VVFAGNGIGSGPAKDFALL-TCQNTHTIMII--GTGIIWAAYLAGGDTIYLANI-TLPDSP 304
CC | | | | | : : : : : | | | | | : : : : : | | | | | :
```

```
Db 278 KYFVS---ENTPQDDLAYSHYSCDATLITAPSTFGWVLGYLSKGOAVYQDIRSTNDVN 334
QY 305 FLK-VFKPEAAFLPEWVGIPAD 325
Db 335 YKKGVLDPDFDFVPSWTSIMLD 356
RESULT 13
VGLD_HSVK
ID VGLD_HSVK STANDARD; PRT; 442 AA.
AC P22484;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein D precursor (glycoprotein 17/18).
GN GD OR GP17/18 OR 72.
OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082407; PubMed=1845821;
RA Flowers C.C., Eastman E.M., O'Callaghan D.J.;
RT "Sequence analysis of a glycoprotein D gene homolog within the unique
RT short segment of the EHV-1 genome.";
RL Virology 180:175-184(1991).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=92263758; PubMed=13156673;
RA Colle C.F., II, Flowers C.C., O'Callaghan D.J.;
RT "Open reading frames encoding a protein-kinase, homolog of
RT glycoprotein gX of pseudorabies virus, and a novel glycoprotein map
RT within the unique short segment of equine herpesvirus type 1.";
RL Virology 188:545-557(1992).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M62923; AAA45081.1; -
CC DR EMBL: M86931; -; NOT_ANNOTATED_CDS.
CC DR EMBL: M87497; AAA45073.1; ALT_INIT.
CC PIR; A38518; VGBEPA.
CC InterPro; IPR002896; Herpes_glycop_D.
CC Pfam; PF01537; Herpes_glycop_D; 1.
CC Glycoprotein; Signal; Transmembrane.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 442 GLYCOPROTEIN D.
CC DOMAIN 20 405 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 406 422 POTENTIAL.
CC DOMAIN 423 442 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 442 AA; 49908 MW; 323CDDCA9C9762F05 CRC64;
CC
CC Query Match 4.7%; Score 91; DB 1; Length 442;
CC Best Local Similarity 22.7%; Pred. No. 3;
CC Matches 54; Conservative 31; Mismatches 103; Indels 50; Gaps 10;
```

```
QY 6 LOQRIVKLO-----PLSEKELPMTQSSGNETSPEMKRDSEQHNGELRGFTI 55
| | | | | : : : : : | | | | | : : : : : | | | | | :
Db 234 LYRVLEIDGRIRYDFSVTIPSEPCPIAFELNFGN---PDRCKTPEQYSRGF---VFTR 287
| | | | | : : : : : | | | | | : : : : : | | | | | :
QY 56 NSIGRLGNMGEXATLALARMGR--LATIPASMHNALAPIFRISLPVLHSDTAKKIPW 113
| | | | | : : : : : | | | | | : : : : : | | | | | :
```

Db 288 RFLGEFNPGEHMTWKWFVYDGNL-----PVQFYEAQA-----FARVPP 331

QY 114 QNYHLNDMEERYH-----IPGHVFRFTGPGSWTFYHHLRPEILKEFTLHDHVRDEAQ 168

Db 332 DNHFGDSVSEITQNTDKPGQADPKPQPKWPSIKHLVPEL-----DEVDEVI 384

QY 169 AFLRGLRVGQSSTFFVGVHVRGDDYVHVMPNWKGVVADRGLYKLEKALDMFRARYSSP 226

Db 385 PVTKPKTSKN-STFVGISVGLGIAGLVGLVGLYVCLRR--KKEKLVCTERLDSF 438

RESULT 14

ID FNA3_PORGI STANDARD; PRT; 353 AA.

AC Q51826; 10

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Major fimbrial subunit protein, type III precursor (Fimbrillin)

DE (Fimbrillin).

GN FIMA.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;

OC Porphyromonadaceae; Porphyromonas.

OX NCBI_TaxID=837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=6/26;

RX MEDLINE=94071950; PubMed=7902712;

RA Fujiwara T., Morishima S., Takahashi I., Hamada S.;

RT "Molecular cloning and sequencing of the fimbrial gene of

RT Porphyromonas gingivalis strains and characterization of recombinant

RT proteins."

RL Biochem. Biophys. Res. Commun. 197:241-247(1993).

RN [2]

RP FUNCTION, AND CLASSIFICATION INTO TYPES.

RX MEDLINE=21614934; PubMed=11748193;

RA Nakagawa I., Amano A., Kubonawa M., Nakamura T., Kawabata S.,

RA Hamada S.;

RT "Functional differences among FimA variants of Porphyromonas

RT gingivalis and their effects on adhesion to and invasion of human

RT epithelial cells."

RL Infect. Immun. 70:277-285(2002).

CC -!- FUNCTION: Fimbrillin is the structural subunit of the fimbriae,

CC that are filamentous appendages on the cell surface. Fimbriae of

CC P. gingivalis are recognized as a major virulence factor as they

CC mediate cell adhesion and play an important role in invasion of

CC periodontal tissues.

CC -!- SUBCELLULAR LOCATION: Fimbria.

CC -!- SIMILARITY: BELONGS TO THE P.GINGIVALIS FIMBRILLIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; D17801; BAA04627.1; -

FW Fimbria; Virulence.

KW PROPEP 1 10 BY SIMILARITY.

FT CHAIN 11 353 MAJOR FIMBRILLIN SUBUNIT PROTEIN, TYPE III.

SQ SEQUENCE 353 AA; 38024 MW; 7FBE4FBF42FEA2AB CRC64;

Query Match 4.6%; Score 89; DB 1; Length 353;

Best Local Similarity 22.1%; Pred. No. 3.3;

Matches 64; Conservative 38; Mismatches 110; Indels 78; Gaps 12;

QY 23 MTIOMSSCNTSPERRDSEQ-----HGNGELRGMTTINSIGRLNQMG---EYATLFA 73

Db 80 LTTLETGNOAAGLIMTAEPVEVTLVAGN-----NYYGVDGSGGQNGISQGTPL 130

QY 74 LARNGRIAP--IPASMHNLAPITRISLPVLHSDTAKK----- 110

Db 131 IKRVHARIAFTKIEVTSQSYANKYNFAPENIYALVAKKSNLFGASLANSDDAYLTGSL 190

QY 111 -----IPWNYHLNDMEERYH-----PGHVFRTGPGSWTFYHH--LRPEIL- 154

Db 191 ITFNGAYSPANYTHVDWLGRDYTGIGATVNTPKGF-----YVLESTYAGNAGLRPTILC 245

QY 155 --KEFTLHDHV-----REEAQFLRGLRVGQSSTFF--VGHVHRRGDDYVHVMPNWKGV 206

Db 246 VKGLTKHDTALSSSEEMTAFAFNAGWIVANNDPTTYYPVLNPFESNNYTYTGEAVEGKI 305

QY 207 ADRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRNINASRGDVVVFAGNGI 256

Db 306 VRNHKFDINLTITGPGTNNPENPIT-----ESANLNVNVCYAAKGV 347

RESULT 15

ID CDH_PHACH STANDARD; PRT; 773 AA.

AC Q01738; Q00047;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cellobiose dehydrogenase precursor (EC 1.1.5.1) (CDH) (Cellobiose-

DE quinone oxidoreductase).

GN CDH-1 AND CDH-2.

OS Phanerochaete chrysosporium.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Corticiaceae; Phanerochaete.

OX NCBI_TaxID=5306;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97077226; PubMed=8919793;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cloning of a cDNA encoding cellobiose dehydrogenase, a

RT hemoflavoenzyme from Phanerochaete chrysosporium."

RL Appl. Environ. Microbiol. 62:1329-1335(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97176414; PubMed=9023960;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cellobiose dehydrogenase from Phanerochaete chrysosporium is encoded

RT by two allelic variants."

RL Appl. Environ. Microbiol. 63:796-799(1997).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.

RX MEDLINE=20139694; PubMed=10673428;

RA Hallberg B.M., Bergfors T., Boeckro K., Pettersson G., Henriksson G.,

RA Dyne C.;

RT "A new scaffold for binding haem in the cytochrome domain of the

RT extracellular flavocytochrome cellobiose dehydrogenase."

RL Structure 8:79-88(2000).

CC -!- FUNCTION: DEGRADATES BOTH LIGNIN AND CELLULOSE. OXIDIZES CELLOBIOSE

CC TO CELLOBIONOLACTONE.

CC -!- CATALYTIC ACTIVITY: Cellobiose + a quinone = cellobiono-1,5-

CC lactone + a phenol.

CC -!- COFACTOR: ONE FAD AND ONE HEME B.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMC

CC OXIDOREDUCTASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; U46081; AAC49277.1; -

```
DR EMBL: U65888; AAB61455.1; -.
DR EMBL: U50409; AB92262.1; -.
DR PDB: 1D7B; 18-OCT-99.
DR PDB: 1D7C; 18-OCT-99.
DR PDB: 1D7D; 18-OCT-99.
DR InterPro: IPR001172; GMC_oxred.
DR Pfam: PF00732; GMC_oxred; 1.
DR PRINTS: PR00411; FNDKDTASEL.
DR PROSITE: PS00623; GMC_OXRED_1; 1.
DR PROSITE: PS00624; GMC_OXRED_2; 1.
KW Cellulose degradation; Oxidoreductase; FAD; Flavoprotein; Heme;
KW Multigene family; Signal; 3D-structure.
FT SIGNAL 1 18
FT CHAIN 19 773 CELLOBIOSE DEHYDROGENASE.
FT DOMAIN 19 208 HEME DOMAIN.
FT DOMAIN 235 773 OXIDOREDUCTASE.
FT BINDING 83 83 HEME LIGAND.
FT BINDING 181 181 HEME LIGAND.
FT NP_BIND 236 265 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 773 AA; 82007 MW; 54F721E779AA4D7B CRC64;

Query Match      4.6%; Score 89; DS 1; Length 773;
Best Local Similarity 22.0%; Pred. No. 8.6;
Matches 49; Conservative 29; Mismatches 91; Indels 54; Gaps 11;

QY 131 GHVYFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFLEG-----LRVNGSQPS 182
Db 337 GDFSSVGVWPSWTHAPYTSKLSRLPSTDHPSTDGQRYLEQSFNVVSQLKGGYNQA 396
QY 183 TFVGVHVRRGDYVHVMPNVKGVVADRGYLEKALDMPFARYSSPVFVTSNGMAWCRENI 242
Db 397 T-----INDNPNYKDHPF---GY--SAEDFLNGKRAQGVATYLTOTALA--RPNF 438
QY 243 -----NASRGDVFVAGNGIEGSPAKD-----FALLTOCNHTIMTIGTEG----IWAAYL 287
Db 439 TEKTNVWVSNVWRNGSILGVQINDPTLGNQFIPVTPKGRVILSAGAFGTSRILFOSGI 498
QY 288 AGGDTIYLANVTLPSDFLKVFPFAFLP---EWVGIPADLS 327
Db 499 GPTDMI-----QTVQSNP-----TAAALPPQWQWNLNLPVGMN 531
```

Search completed: May 27, 2003, 15:07:13
Job time : 11.478 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 ; Search time 40.5157 Seconds
(without alignments)
1932.533 Million cell updates/sec

Title: US-10-040-863-8

Perfect score: 2032

Sequence: 1 MASAAQVFPSPPLAHFLIFV.....RSHFLKAKGVTCYVAGRAF 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_cheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2032	100.0	380	11 Q9JK44	Q9JK44 rattus norv
2	1896	93.3	353	11 O55025	O55025 rattus norv
3	1883	92.7	354	11 O35087	O35087 rattus norv
4	1878	92.4	354	11 Q38275	Q38275 rattus norv
5	1674.5	82.4	347	11 Q9UL27	Q9UL27 mus musculus
6	1670.5	82.2	347	11 Q70504	Q70504 mus musculus
7	1547.5	76.2	321	11 Q920W2	Q920W2 mus spicile
8	1547.5	76.2	321	11 Q91VF0	Q91VF0 mus musculus
9	1532.5	75.4	321	11 Q920W3	Q920W3 mus musculus
10	1499.5	73.8	343	6 Q77486	Q77486 gorilla gor
11	1495.5	73.6	343	6 Q77485	Q77485 pan troglod
12	1495.5	73.6	343	6 Q9TUD3	Q9TUD3 gorilla gor
13	1488.5	73.3	338	6 Q9TUD5	Q9TUD5 pan troglod
14	1488.5	73.3	347	6 Q29505	Q29505 cryotolagus
15	1487.5	73.2	343	6 Q77487	Q77487 pongo pygma
16	1478.5	72.8	343	6 Q9TTC7	Q9TTC7 hylobates l

ALIGNMENTS

RESULT 1

Q9JK44 ID Q9JK44 PRELIMINARY; PRT; 380 AA.

AC Q9JK44;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Alpha 1-2 fucosyltransferase.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FISCHER;

RA Sherwood A.L., Stroud M.R., Levery S.B., Holmes F.H.;

RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-

RT fucose:GMI alpha 1-2 fucosyltransferase is required for optimum enzyme

RT activity and interaction with lipids."

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF264005; AAF72200.1;

DR InterPro; IPR002516; GT_11.

DR Pfam; PF01531; Glyco.transf_11; 1.

KW Glycosyltransferase; Transferase.

SQ SEQUENCE 380 AA; 42742 MW; 920592C96689C05C CRC64;

Query Match 100.0%; Score 2032; DB 11; Length 380;

Best Local Similarity 100.0%; Pred. No. 1.8e-176;

Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAAQVFPSPPLAHFLIFVVTSTIIHLQQRIVKQLPSEKELPMTTQMSGNTSEPMR 60

Db 1 MASAAQVFPSPPLAHFLIFVVTSTIIHLQQRIVKQLPSEKELPMTTQMSGNTSEPMR 60

QY 61 RDSFQHGNGELRGMPFINSIGRLGNQMGAYATLFPALRMNGRLAFIPASMHNALAPIFRI 120

Db 61 RDSFQHGNGELRGMPFINSIGRLGNQMGAYATLFPALRMNGRLAFIPASMHNALAPIFRI 120

QY 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180

Db 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180

QY 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180

Db 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180

QY 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180

Db 121 SLPVLHSDTAKKIPQWYHLNDNMWEERYRHPGHFVRFYHHLRPEILKEFT 180


```

Db 121 SLPVLHSDTAKKIPWNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEFT 180
QY 181 LHDHVREEAQAFGLRLVNGSOPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
Db 181 LHDHVREEAQAFGLRLVNGSOPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHITMTIG 300
Db 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHITMTIG 300
QY 301 TFGIWAAYLAGDGTIYLANITPLDPSFPLKVPKPEAFLEPWGIPADLSPLKALTPAC 360
Db 301 TFGIWAAYLAGDGTIYLANITPLDPSFPLKVPKPEAFLEPWGIPADLSPLKALTPAC 360
QY 361 RSHFHLKAGVTCYVAGRAF 380
Db 361 RSHFHLKAGVTCYVAGRAF 380

RESULT 2
Q5025
AC O55025 PRELIMINARY; PRT; 353 AA.
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BUFFALO RAT;
RX MEDLINE=98342056; PubMed=9675030;
RA Sherwood A.L., Holmes E.H.;
RT "Cloning and expression of the catalytic domain from rat hepatoma H35
RT cell GGP-fucose: GMI alpha 1-->2fucosyltransferase, an enzyme which is
RT activated during early stages of chemical carcinogenesis in rat
RT liver.";
RL Arch. Biochem. Biophys. 355:215-221(1998).
DR EMBL: AF042743; AAC14695.1;
DR InterPro: IPR002516; GT.11.
DR Pfam: PF01531; Glyco_transf.11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 353 AA; 39738 MW; 609D64EB22C9585 CRC64;

Query Match 93.3%; Score 1896; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.8e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQORIVKLOPLSEKELPMTOMSSGNTSPENRSEORNGELRGMTFINSIGRLGNOM 87
Db 1 LQORIVKLOPLSEKELPMTOMSSGNTSPENRSEORNGELRGMTFINSIGRLGNOM 60
QY 88 GYATLALARNGRALFIPASMHNALAPIFRISLFLVLSHSDTAKKIPWNYHLNDWMEER 147
Db 61 GYATLALARNGRALFIPASMHNALAPIFRISLFLVLSHSDTAKKIPWNYHLNDWMEER 120
QY 148 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRLVNGSOPSTFV 207
Db 121 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFGLRLVNGSOPSTFV 180
QY 208 GYHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAMCARENINAS 267
Db 181 GYHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPVVVTSGMAMCARENINAS 240
QY 268 RGDVVFAGNGIEGSPAKDFALLTQCNHITMTIGTFGIWAAYLAGDGTIYLANITPLDPSF 327
Db 241 RGDVVFAGNGIEGSPAKDFALLTQCNHITMTIGTFGIWAAYLAGDGTIYLANITPLDPSF 300

```

```

QY 328 LKVFKEPAFLPEWGIADLSPLKALTPACPRSHFHLKAKGVTCTYVAGRAF 380
Db 301 LKVFKEPAFLPEWGIADLSPLKALTPACPRSHFHLKAKGVTCTYVAGRAF 353

RESULT 3
Q55087
ID O35087 PRELIMINARY; PRT; 354 AA.
AC O35087;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON CANCER;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RT fucosyltransferase genes.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AB006138; BAA21742.1;
DR InterPro: IPR002516; GT.11.
DR Pfam: PF01531; Glyco_transf.11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 86364444889215BA1 CRC64;

Query Match 92.7%; Score 1883; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.8e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAQVPFSEPLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTOMSSGNTSPENR 60
Db 1 MASAQVPFSEPLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTOMSSGNTSPENR 60
QY 61 RSEQHNGELRGMTFINSIGRLGNOMGEYATLFLALARNGRALFIPASMHNALAPIFR 120
Db 61 RSEQHNGELRGMTFINSIGRLGNOMGEYATLFLALARNGRALFIPASMHNALAPIFR 120
QY 121 SLPVLHSDTAKKIPWNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEFT 180
Db 121 SLPVLHSDTAKKIPWNYHLNDWMEERYHIFGHFVRFYPCSWTFYHHLRPEILKEFT 180
QY 181 LHDHVREEAQAFGLRLVNGSOPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
Db 181 LHDHVREEAQAFGLRLVNGSOPSTFVGHVHVRGDYVHVMPNWKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHITMTIG 300
Db 241 FRARYSSPVVVTSGMAMCARENINASRGDVVFAGNGIEGSPAKDFALLTQCNHITMTIG 300
QY 301 TFGIWAAYLAGDGTIYLANITPLDPSFPLKVPKPEAFLEPWGIPADLSPLK 353
Db 301 TFGIWAAYLAGDGTIYLANITPLDPSFPLKVPKPEAFLEPWGIPADLSPLK 353

RESULT 4
Q9R275
ID Q9R275 PRELIMINARY; PRT; 354 AA.
AC Q9R275;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

```



```

[1]
SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RA MEDLINE=21099374; PubMed=11179967;
RX Bureau V., Marionneau S., Cailleteau-Thomas A., Le Mouillac-Vaidye B.,
RA Liehr T., Le Pendu J.;
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL: AF131238; AAD24469.1; -
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39983 MW; 123E8C837928559E CRC64;

Query Match 92.4%; Score 1878; DB 11; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.7e-162;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKELPMTTOMSSGNTSPENR 60
DB 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKELPMTTOMSSGNTSPENR 60
QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
DB 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
QY 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 180
DB 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 180
QY 151 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVVADRGYLEKALDM 240
DB 151 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVVADRGYLEKALDM 240
QY 241 FRARYSSPVFVVTNSGMWACRENINASRGDVVFAGNGTEGSPAKDFALLTCQNHHTIMTIG 300
DB 241 FRARYSSPVFVVTNSGMWACRENINASRGDVVFAGNGTEGSPAKDFALLTCQNHHTIMTIG 300
QY 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVKPEAAFLPEWVGIPADLSPLIK 353
DB 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVKPEAAFLPEWVGIPADLSPLIK 353

RESULT 5
Q9JL27 PRELIMINARY; PRT; 347 AA.
ID Q9JL27
AC Q9JL27;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)Fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL: AF214656; AAF45146.1; -
DR MGB: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046CC8E79 CRC64;

```

```

Query Match 82.4%; Score 1674.5; DB 11; Length 347;
Best Local Similarity 89.8%; Pred. No. 5.3e-144;
Matches 317; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

QY 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKELPMTTOMSSGNTSPENR 60
DB 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKEL-QAVQMSSPNAARTDMQ 59
QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
DB 60 Q-----SAKLOGIFTINSIGRLGNOMGEYATLFLALARMNGRLAFIPESMHNALAPIERI 113
QY 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFT 180
DB 114 SLPVLHSDTARRIPWQNYHLNDWMEERYRHIPGVFRFTGYPCSWTFYHHLRPEILKEFT 173
QY 151 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVVADRGYLEKALDM 240
DB 174 LHDHVREAAQAFRLGLRVNGSQPSTFVGHVHVRGDIYVHVMNVKGVVADRGYLEKALOR 233
QY 241 FRARYSSPVFVVTNSGMWACRENINASRGDVVFAGNGTEGSPAKDFALLTCQNHHTIMTIG 300
DB 234 FRARYSSPVFVVTNSGMWACRENINATSLGDVVFAGNGTEGSPAKDFALLTCQNHHTIMTIG 293
QY 301 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVKPEAAFLPEWVGIPADLSPLIK 353
DB 294 TFGIWAAYLAGDGTIYLANYLTPDSPFLKVKPEAAFLPEWVGIPADLSPLIK 346

RESULT 6
O70504 PRELIMINARY; PRT; 347 AA.
ID O70504
AC O70504;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RT "GDP-fucose: beta-galactoside alphas,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide.";
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL: AF084792; AAC16887.1; -
DR MGB: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;

Query Match 82.2%; Score 1670.5; DB 11; Length 347;
Best Local Similarity 89.5%; Pred. No. 1.2e-143;
Matches 316; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

QY 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKELPMTTOMSSGNTSPENR 60
DB 1 MASAOVPSFPLAHLFIHFVFTSTIIHQORIVKLOPLSEKEL-QAVQMSSPNAARTDMQ 59
QY 61 RDSQHGNGELRGMTINSIGRLGNOMGEYATLFLALARMNGRLAFIPASMHNALAPIERI 120
DB 60 Q-----SAKLOGIFTINSIGRLGNOMGEYATLFLALARMNGRLAFIPESMHNALAPIERI 113

```



```

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE GpL-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase (EC
DE 2.4.1.69) (Fragment).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NJL/MSF;
RA Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039122; BAF68646.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1 1
FT NON_TER 321 321
SQ SEQUENCE 321 AA; 36464 MW; FF5304CD150F774A CRC64;

Query Match 75.4%; Score 1532.5; DB 11; Length 321;
Best Local Similarity 88.4%; Pred. No. 4e-131;
Matches 290; Conservative 13; Mismatches 18; Indels 7; Gaps 2;

QY 15 FLIFVFVTSIIHQRIKVLQPLSEKELPMTQMSSGNTSPESPMRDSQHGNGELRGM 74
DB 1 FLIFVFVTSIIHQRIKVLQPLSEKELPMTQMSSGNTSPESPMRDSQHGNGELRGM 74
QY 75 FTINSIGRLGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIP 134
DB 54 FTINSIGRLGNQGEYATLALARMNGRLAFIPESMHNALAPIFRISLPVLHSDTARIP 113
QY 135 WQNYHLNDWMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAF 194
DB 114 WQNYHLNDWMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAF 173
QY 195 GLRVNGSQPTFVGVHVVRGDDYVHVMPKVMGVVADRGYLEKALDMFRARYSSPVFVVT 254
DB 174 GLRVNGSQPTFVGVHVVRGDDYVHVMPKVMGVVADRGYLEKALDMFRARYSSPVFVVT 233
QY 255 NGKAWCRENINASRGDVFVAGNTEGSPAKDFALLTCNHTIMTIGTFGWAAYLAGGDT 314
DB 234 NGKAWCRENINTSGDVFVAGNTEGSPAKDFALLTCNHTIMTIGTFGWAAYLAGGDT 293
QY 315 IYLANYTLPDPSPLKVPKPAAPLPEWV 342
DB 294 IYLANYTLPDPSPLKVPKPAAPLPEWV 321

RESULT 10
O77486 PRELIMINARY; PRT; 343 AA.
AC O77486;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL: AB015635; BAA31128.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FBC CRC64;

Query Match 73.8%; Score 1499.5; DB 6; Length 343;
Best Local Similarity 78.5%; Pred. No. 4.4e-128;
Matches 277; Conservative 29; Mismatches 36; Indels 11; Gaps 2;

QY 1 MASQVFPSPFLAHLFLIFVFTSTIIHQRIKVLQPLSEKELPMTQMSSGNTSPESPMR 60
DB 1 MLVYQMPFSPFMAHFLIFVFTSTIFVHQRLAKIQAM--WELEFV-----QIPVLA 49
QY 61 RDSQHGNGELRGMFTINSIGRLGNQGEYATLALARMNGRLAFIPASMHNALAPIFR 120
DB 50 STSKALGPSQLRGWWTINAIGRLGNQGEYATLALAKMNGRPAPFPAQMSTLAPIFR 109
QY 121 SLPLVHSDTAKKIPQNYHLNDWMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEFT 180
DB 110 TLPVLHSATASRIQWNYHLNDWMEERYRHIPGHEVFTGYPCSWTFYHHLRQELQDEFT 169
QY 181 LHDHVREREAQAFGLRVNGSQPTFVGVHVVRGDDYVHVMPKVMGVVADRGYLEKALDM 240
DB 170 LHDHVREREAQAFGLRVNGSQPTFVGVHVVRGDDYVHVMPKVMGVVADRGYLEKALDM 229
QY 241 FRARYSSPVFVVTSGKAWCRENINASRGDVFVAGNTEGSPAKDFALLTCNHTIMTIG 300
DB 230 FRARYSSPVFVVTSGKAWCRENINTSGDVFVAGDTEGSPAKDFALLTCNHTIMTIG 289
QY 301 TFGWAAYLAGGDTIYLANYTLPDPSPLKVPKPAAPLPEWVGIPADLSPLIK 353
DB 290 TFGWAAYLAGGDTIYLANYTLPDPSPLKVPKPAAPLPEWVGIPADLSPLIK 342

RESULT 11
O77485 PRELIMINARY; PRT; 343 AA.
AC O77485;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL: AB015634; BAA31127.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match. 73.6%; Score 1495.5; DB 6; Length 343;

```

```
Best Local Similarity 78.2%; Pred. No. 1e-127;
Matches 276; Conservative 31; Mismatches 35; Indels 11; Gaps 2;

QY 1 MASQVFPSPFLAHFLFVFTVSTIIHLQORIVKQLPLSEKELPMTTQMSSGNTSEPMR 60
Db 1 MLVVQMPFSPFAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
QY 61 RDSQHGNGELRGMTINSTIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRI 120
Db 50 STSKALGPSOLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMHS TLAPFIRI 109
QY 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 180
Db 110 TLPVLHSATASRIPWQNYHLNDWMEERYRHIPGEYVFTGYPCSWTFYHHLRQEILOEFT 169
QY 181 LHDHVRREAQAFRLGLRVNGSOPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
Db 170 LHDHVRREAQKFLRLGLQVNGSRPGTFFGVHVRRGDYVHVMPKVGWVADRRYLOQALDW 229
QY 241 FRARYSSPFFVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCNHTIMTIG 300
Db 230 FRARYSSPIFFVVTSGMAWCRENIDTSGDVVFAGDIEGSPAKDFALLTCNHTIMTIG 289
QY 301 TFGIWAAYLAGGDTIYLANTYLTPDSPPLKVFKEPAEALPEWVGIPADLSPLK 353
Db 290 TFGIWAAYLTGGDTIYLANTYLTPDSPPLKIFKPEAAFLPEWMTGIAADLSPLK 342

RESULT 12
Q9TUD3
ID Q9TUD3 PRELIMINARY; PRT; 343 AA.
AC Q9TUD3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha (1.2) fucosyl transferase.
GN FUT2
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080606; AAF14068.1;
DR InterPro; IPR002516; GT.11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AE8C6C CRC64;

Query Match 73.6%; Score 1495.5; DB 6; Length 343;
Best Local Similarity 78.2%; Pred. No. 1e-127;
Matches 276; Conservative 30; Mismatches 36; Indels 11; Gaps 2;

QY 1 MASQVFPSPFLAHFLFVFTVSTIIHLQORIVKQLPLSEKELPMTTQMSSGNTSEPMR 60
Db 1 MLVVQMPFSPFAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
QY 61 RDSQHGNGELRGMTINSTIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRI 120
Db 50 STSKALGPSOLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMHS TLAPFIRI 109
QY 121 SLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT 180
Db 110 TLPVLHSATASRIPWQNYHLNDWMEERYRHIPGEYVFTGYPCSWTFYHHLRQEILOEFT 169
```

```
QY 181 LHDHVRREAQAFRLGLRVNGSOPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDM 240
Db 170 LHDHVRREAQKFLRLGLQVNGSRPGTFFGVHVRRGDYVHVMPKVGWVADRRYLOQALDW 229
QY 241 FRARYSSPFFVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCNHTIMTIG 300
Db 230 FRARYSSPIFFVVTSGMAWCRENIDTSGDVVFAGDIEGSPAKDFALLTCNHTIMTIG 289
QY 301 TFGIWAAYLAGGDTIYLANTYLTPDSPPLKVFKEPAEALPEWVGIPADLSPLK 353
Db 290 TFGIWAAYLTGGDTIYLANTYLTPDSPPLKIFKPEAAFLPEWMTGIAADLSPLK 342

RESULT 13
Q9TUD5
ID Q9TUD5 PRELIMINARY; PRT; 338 AA.
AC Q9TUD5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha (1.2) fucosyl transferase.
GN FUT2
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHI6;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Desplau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080604; AAF14066.1;
DR InterPro; IPR002516; GT.11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 338 AA; 38428 MW; E7FABA0FF1BC95F9 CRC64;

Query Match 73.3%; Score 1488.5; DB 6; Length 338;
Best Local Similarity 78.7%; Pred. No. 4.3e-127;
Matches 274; Conservative 31; Mismatches 32; Indels 11; Gaps 2;

QY 6 VPSFPLAHFLIFVFTVSTIIHLQORIVKLOPLSEKELPMTTQMSSGNTSEPMRRDSEQ 65
Db 1 MPFSFPVAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA SKA 49
QY 66 HGNGELRGMTINSTIGRLGNQMGEXATLAFALARMNGRLAFIPASMHNALAFIRISLPVL 125
Db 50 LGPSQLRGMTINAGRLGNQMGEXATLYALAKMNGRPAFIPAQMHS TLAPFIRI LPVL 109
QY 126 HSDTAKKIPWQNYHLNDWMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKEFT LHDV 185
Db 110 HSATASRIPWQNYHLNDWMEERYRHIPGEYVFTGYPCSWTFYHHLRQEILOEFT LHDV 169
QY 186 REBAQAFRLGLRVNGSOPSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFEARY 245
Db 170 REBAQKFLRLGLQVNGSRPGTFFGVHVRRGDYVHVMPKVGWVADRRYLOQALDWPRARY 229
QY 246 SSFVVTSGMAWCRENINASRGDVVFAGNGIEGSPAKDFALLTCNHTIMTIGTIGFIW 305
Db 230 SSFVVTSGMAWCRENIDTSGDVVFAGDIEGSPAKDFALLTCNHTIMTIGTIGFIW 289
QY 306 AAYLAGGDTIYLANTYLTPDSPPLKVFKEPAEALPEWVGIPADLSPLK 353
Db 290 AAYLTGGDTIYLANTYLTPDSPPLKIFKPEAAFLPEWMTGIAADLSPLK 337

RESULT 14
Q29505
```

Search completed: May 27, 2003, 15:10:44
Job time : 43.5157 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 : Search time 38.17 seconds
(without alignments)
1932.533 Million cell updates/sec

Title: US-10-040-863-8_COPY_23_380

Perfect score: 1921

Sequence: 1 STIIHQRIKVLQPLSEKE.....RSHFLKAKGVTCYVYAGRAF 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921	100.0	380	11	Q9JK44
2	1896	98.7	353	11	O55025
3	1772	92.2	354	11	O35087
4	1767	92.0	354	11	O9R275
5	1563.5	81.4	347	11	O9JL27
6	1559.5	81.2	347	11	O70504
7	1508.5	78.5	321	11	O920W2
8	1508.5	78.5	321	11	O9IVF0
9	1493.5	77.7	321	11	O920W3
10	1421.5	74.0	343	6	O77486
11	1418.5	73.8	338	6	O9TUD5
12	1418.5	73.8	343	6	O77485
13	1417.5	73.8	343	6	O9TUD3
14	1412.5	73.5	347	6	O29505
15	1410.5	73.4	343	6	O77487
16	1403.5	73.1	332	4	O99450

17	1403.5	73.1	343	6	O77712
18	1400.5	72.9	343	6	O9TTC7
19	1354.5	70.5	368	11	O920T3
20	1344.5	70.0	346	4	Q14338
21	1343	69.9	299	11	O9WUE6
22	1342.5	69.9	368	11	O920T2
23	1338.5	69.7	368	11	O920T4
24	1338.5	69.7	368	11	O91V73
25	1319	68.7	273	6	O9TSL5
26	1307	68.0	273	6	O9TTC9
27	1262	65.7	273	6	O9TUD0
28	1237.5	64.4	348	6	O9TUD1
29	1215.5	63.3	332	6	O9N266
30	1214	63.2	368	6	O9TYY3
31	1109	57.7	377	11	O9JL28
32	1109	57.7	377	11	P97327
33	1099	57.2	377	11	O920W5
34	1099	57.2	377	11	O920W4
35	1086.5	56.6	360	6	O9TTY7
36	1051.5	54.7	366	6	O9TUE6
37	1051.5	54.7	366	6	O9TUD4
38	1048.5	54.6	365	6	O9TUE7
39	1048.5	54.6	366	6	O9TUD2
40	1047.5	54.5	353	6	O9N270
41	1047.5	54.5	365	4	O8XRB6
42	1047.5	54.5	366	6	O9R267
43	1046.5	54.5	365	4	O14505
44	1043.5	54.3	366	6	O9TUD6
45	1042.5	54.3	365	4	O14506

ALIGNMENTS

RESULT 1

O9JK44 PRELIMINARY; PRT; 380 AA.
AC O9JK44;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha 1-2 fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomii;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FISCHER;
RA Sherwood A.L.; Stroud M.R., Levery S.B., Holmes E.H.;
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-fucose:GMI alpha 1-2 fucosyltransferase is required for optimum enzyme activity and interaction with lipids."
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF264005; AAF72200.1;
DR InterPro; IPR002516; GI_11;
DR Pfam; PF01531; Glyco_transf_11; 1.
DR Glycosyltransferase; Transferase.
KW SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;
Query Match 100.0%; Score 1921; DB 11; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.2e-166;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIIHQRIKVLQPLSEKELPMTTQSSGNTESPMMRDSPOHNGELRGMTINSIGR 60

Db 23 STIIHQRIKVLQPLSEKELPMTTQSSGNTESPMMRDSPOHNGELRGMTINSIGR 82

QY 61 LGNQMGYATLFALARMNGRLAFTIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLND 120

Db 83 LGNQMGYATLFALARMNGRLAFTIPASMHNALAPIFRISLPVLHSDTAKKIPQWYHLND 142

QY 121 WMEERYHRTPGHEVRFYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180

|||||
Db 143 WMEERYRHIPGHEVFTGPGSWTFYHHLRPEILKEFTLHDHVREAAQAFGLRVNGSQ 202
QY 161 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 203 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDGVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 300
Db 263 NINASRGDGVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 322
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 358
Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 380

RESULT 2
O55025 PRELIMINARY: PRT: 353 AA.
AC O55025;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase (fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=BUFFALO RAT;
RX MEDLINE=98342056; PubMed=9675030;
RA "Shrood A.L., Holmes E.H.;
RT "Cloning and expression of the catalytic domain from rat hepatoma H35
RT cell GbP-fucose; GMI alpha 1-->2fucosyltransferase, an enzyme which is
RT activated during early stages of chemical carcinogenesis in rat
RT liver.";
RL Arch. Biochem. Biophys. 355:215-221(1998).
DR EMBL; AF042743; AAC14695.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco.transf.11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER
SQ SEQUENCE 353 AA; 39738 MW; 609D64EB222C9585 CRC64;

Query Match 98.7%; Score 1896; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.1e-164;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LOORIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGRLGNOM 65
Db 1 LOORIVKLOPSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGRLGNOM 60
QY 66 GYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHSDTAKKIPQWNYHLND 125
Db 61 GYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHSDTAKKIPQWNYHLND 120
QY 126 YRHIPGHEVFTGPGSWTFYHHLRPEILKEFTLHDHVREAAQAFGLRVNGSQPSTFV 185
Db 121 YRHIPGHEVFTGPGSWTFYHHLRPEILKEFTLHDHVREAAQAFGLRVNGSQPSTFV 180
QY 186 GVHVRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCREINAS 245
Db 181 GVHVRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCREINAS 240
QY 246 RGVVVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 305
Db 241 RGVVVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 300
QY 306 LKVEKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 358
Db 301 LKVEKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFHLKAKGVTCYVAGRAF 353

RESULT 3
O35087 PRELIMINARY: PRT: 354 AA.
AC O35087;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
GN FIG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=COLON CANCER;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RL fucosyltransferase genes.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006138; BAA21742.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco.transf.11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 8636444488215BA1 CRC64;

Query Match 92.2%; Score 1772; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 4.2e-153;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIIHLQRIKVLQPLSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGR 60
Db 23 STIIHLQRIKVLQPLSEKELPMTQSSGNTSEPMRRDSEHGNGELRGMTINSIGR 82
QY 61 LGNQMEYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHSDTAKKIPQWNYHLND 120
Db 83 LGNQMEYATLALARMNGRLAFIPASMHNALAPIRISLPVLSHSDTAKKIPQWNYHLND 142
QY 121 WMEERYRHIPGHEVFTGPGSWTFYHHLRPEILKEFTLHDHVREAAQAFGLRVNGSQ 180
Db 143 WMEERYRHIPGHEVFTGPGSWTFYHHLRPEILKEFTLHDHVREAAQAFGLRVNGSQ 202
QY 181 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 203 PSTFVGHVRRGDYVHVPNVKGVVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDGVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 300
Db 263 NINASRGDGVFAGNIEGSPAKDFALLTQCNHTIMTIGTFGIWAAYLAGGDTIYLANVTL 322
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLLK 331
Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLLK 353

RESULT 4
O9R275 PRELIMINARY: PRT: 354 AA.
AC O9R275;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FIG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=21099374; PubMed=11179967;


```

RA Bureau V., Marionneau S., Caillaud-Thomas A., Le Moullac-Vaidre B.,
RA Liehr T., Le Pendu J.;
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RT alpha-L-fucosyltransferases FTA, FTB and FTC.";
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL: AF131238; AAD24469.1; -.
DR InterPro: IPR002516; GT_11;
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39983 MW; 123B8C8379B8559E CRC64;

Query Match 92.0%; Score 1767; DB 11; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.2e-152;
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 50
Db 23 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 50
QY 61 LGNQGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 83 LGNQGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 142
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 143 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 202
QY 181 PSTFVGVRGDDVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 203 PSTFVGVRGDDVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 262
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTOCNHTTMTTGTGFIWAAYLAGGDTIYLYNTL 300
Db 263 NINASRGDVVFAGNGIEGSPAKDFALLTOCNHTTMTTGTGFIWAAYLAGGDTIYLYNTL 322
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 331
Db 323 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 353

RESULT 5
Q9JL27 ID Q9JL27 PRELIMINARY; PRT; 347 AA.
AC Q9JL27
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=2136545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL: AF214656; AAF45146.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046CC8E79 CRC64;

Query Match 81.4%; Score 1563.5; DB 11; Length 347;
Best Local Similarity 89.1%; Pred. No. 4.1e-134;
Matches 295; Conservative 12; Mismatches 17; Indels 7; Gaps 2;

```

```

QY 1 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 60
Db 23 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 75
QY 61 LGNQGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 76 LGNQGEYATLAFALARMNGRLAFIPESMHNALAFIFRISLPVLHSDTARRIPWQNYHLND 135
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 136 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 195
QY 181 PSTFVGVRGDDVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 196 PSTFVGVRGDDVHVMPNVKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 255
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTOCNHTTMTTGTGFIWAAYLAGGDTIYLYNTL 300
Db 256 NINASRGDVVFAGNGIEGSPAKDFALLTOCNHTTMTTGTGFIWAAYLAGGDTIYLYNTL 315
QY 301 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 331
Db 316 PDSPLKVKFPEAAFLPEWVGIPADLSPLK 346

RESULT 6
Q70504 ID Q70504 PRELIMINARY; PRT; 347 AA.
AC Q70504
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RT "GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide.";
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL: AF064792; AAC16887.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C41179 CRC64;

Query Match 81.2%; Score 1559.5; DB 11; Length 347;
Best Local Similarity 88.8%; Pred. No. 9.4e-134;
Matches 294; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

```

```

QY 1 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 60
Db 23 STIIHQRIIVKLOPLSEKELPMTQMSSGNTSPSPMRDSEHQHNGELRGMTINSIGR 75
QY 61 LGNQGEYATLAFALARMNGRLAFIPASMHNALAFIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 76 LGNQGEYATLAFALARMNGRLAFIPESMHNALAFIFRISLPVLHSDTARRIPWQNYHLND 135
QY 121 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 180
Db 136 WMEERYRHIPGHFVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQ 195

```


QY	181	PSTFVGHVHRGDIYVHMVNKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE	240
!!	!!	!!	!!
Db	196	PSTFVGHVHRSGDIIYHVMFKVKGVADRGYLEKALDRFRARYSSPFVFVTSNGMAWCRE	255
!!	!!	!!	!!
QY	241	NINASRGDVVFAGNGIEGSPAKDFALLTCQNHTIMTICTFGIWAAYLAGGDTIYLANYYTL	300
!!	!!	!!	!!
Db	256	NINTSLGDVVVFAGNGIEGSPAKDFALLTCQNHTIMTICTFGIWAAYLAGGDTIYLANYYTL	315
!!	!!	!!	!!
QY	301	PDSPFLKFVKPEAAFTLPEWVGIPADLSPLLK	331
!!	!!	!!	!!
Db	316	PDSPFLKFIEPAAAFTEPWGIPADLSPLLK	346
!!	!!	!!	!!
RESULT 7			
Q920W2			
ID	Q920W2	PRELIMINARY;	PRT; 321 AA.
AC	Q920W2;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	GDP-l-fucose-beta-D-galactoside 2-alpha-L-fucosyltransferase [EC		
DE	2.4.1.69] (Fragment).		
CN	FUT2.		
OS	Mus spicilegus (Steppe mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
CX	NCBI_TaxID=10103;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ZBN;		
RA	Liu Y., Kitano T., Koide T., Shiroishi T., Moriaki K., Saitou N.;		
RT	"Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of		
RT	Five Mus musculus subspecies."		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB039123; BAB68647.1; .		
DR	InterPro; IPR002516; GT11.		
DR	Dfam; PF01531; Glyco_Transf_11; 1.		
KW	Glycosyltransferase; Transferase.		
FT	NON_TER 1		
FT	NON_TER 321 321		
SQ	SEQUENCE 321 AA; 36560 MW; 5B7D19BA6D4B5394 CRC64;		
Query Match			
Best Local Similarity			
Matches 285; Conservative 11; Mismatches 17; Indels 7; Gaps			
QY	1	STIIHQQRIVKLQPSLEKELPMTQMSSGNTPESMRDRSEQHGNGELRGMFTINSIGR	60
!!	!!	!!	!!
Db	9	STIIHLQQRIVKLQTLSEKEL-QAIOSSLNAARTEMQQ-----SAKIQGFIFNSIGR	61
!!	!!	!!	!!
QY	61	LGNOMGEYATLFALARMGRAPFAPSMMNALAPIFRISLPVLHSDTAKKIPWQNYHLND	120
!!	!!	!!	!!
Db	62	LGNOMGEYATLFALARMGRGLAFIPESHMNLAPIFRISLPVLHSDTARRIPWQNYHLND	121
!!	!!	!!	!!
QY	121	WMEEYRHPIPGHFVRGTGYPCSWTFYHHLRPELLKEFTLDHVREEAQAFRLGRLVNGSQ	180
!!	!!	!!	!!
Db	122	WMEEYRHPIPGQVVRTGYPCSWTFYHHLRPELLKEFTLDHVREEAQAFRLGRLVNGSQ	181
!!	!!	!!	!!
QY	181	PSTFVGHVHRGDIYVHMVNKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE	240
!!	!!	!!	!!
Db	182	PSTFVGHVHRGDIYVHMFKVKGVADRGYLEKALDRFRARYSSPFVFVTSNGMAWCRE	241
!!	!!	!!	!!
QY	241	NINASRGDVVFAGNGIEGSPAKDFALLTCQNHTIMTICTFGIWAAYLAGGDTIYLANYYTL	300
!!	!!	!!	!!
Db	242	NINTSLGDVVVFAGNGIEGSPAKDFALLTCQNHTIMTICTFGIWAAYLAGGDTIYLANYYTL	301
!!	!!	!!	!!
QY	301	PDSPFLKFVKPEAAFTLPEWVGIPADLSPLLK	320
!!	!!	!!	!!
Db	302	PDSPFLKFIEPAAAFTEPWGIPADLSPLLK	321
!!	!!	!!	!!
RESULT 8			
Q91LVFO			

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NUL/MSF;
RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039122; BAB68646.1; -.
DR InterPro; IPR002516; GT_11.
KW Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 321
SQ SEQUENCE 321 AA; 36464 MW; FF5304CD150F774A CRC64;

Query Match 77.7%; Score 1493.5; DB 11; Length 321;
Best Local Similarity 88.1%; Pred. No. 8.7e-128;
Matches 282; Conservative 13; Mismatches 18; Indels 7; Gaps 2;

QY 1 STIIHQORIVKLPSEKELPMTQSSNGTSEPMRDRSEQHNGELRGMTINSIGR 60
Db 9 STIIHQORIVKLPSEKELPMTQSSNGTSEPMRDRSEQHNGELRGMTINSIGR 61
QY 61 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 62 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 121
QY 121 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKETLHDHVREREAQAFGLRVRNGSQ 180
Db 122 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKETLHDHVREREAQAFGLRVRNGSQ 181
QY 181 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 182 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 241
QY 241 NINASRGDVPFAGNIEGSPAKDFALLTQCNHTMTTGTGFINAAYLAGGDTIYLANYYL 300
Db 242 NINSLGDDVVFAGNIEGSPAKDFALLTQCNHTMTTGTGFINAAYLAGGDTIYLANYYL 301
QY 301 PDSPLKLVFKPEAAFLPEW 320
Db 302 PDSPLKLVFKPEAAFLPEW 321

RESULT 10
O77486 PRELIMINARY; PRT; 343 AA.
AC O77486;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";

```

```

J. Biol. Chem. 270:4640-4649(1995).
RL EMBL; AB015635; BAA31128.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 854E28BD74AE8FBC CRC64;

Query Match 74.0%; Score 1421.5; DB 6; Length 343;
Best Local Similarity 79.8%; Pred. No. 3.5e-121;
Matches 264; Conservative 25; Mismatches 31; Indels 11; Gaps 2;

QY 1 STIIHQORIVKLPSEKELPMTQSSNGTSEPMRDRSEQHNGELRGMTINSIGR 60
Db 23 STIFHVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTWTFINALGR 71
QY 61 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 120
Db 72 LGNQMGYATLAFALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLND 131
QY 121 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKETLHDHVREREAQAFGLRVRNGSQ 180
Db 132 WMEERYRHIPGHFVFTGYPCSWTFYHHLRPEILKETLHDHVREREAQAFGLRVRNGSQ 191
QY 181 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 240
Db 192 PSTFVGHVHRRGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMAWCRE 251
QY 241 NINASRGDVPFAGNIEGSPAKDFALLTQCNHTMTTGTGFINAAYLAGGDTIYLANYYL 300
Db 252 NIDTSHGDDVVFAGNIEGSPAKDFALLTQCNHTMTTGTGFINAAYLAGGDTIYLANYYL 311
QY 301 PDSPLKLVFKPEAAFLPEW 331
Db 312 PDSPLKLVFKPEAAFLPEW 342

RESULT 11
O9TUD5 PRELIMINARY; PRT; 338 AA.
AC O9TUD5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2) fucosyl transferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH16.
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT between an intronic Alu-Y element and red cell expression of ABH
RT antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL; AF080604; AAF14056.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase
SQ SEQUENCE 338 AA; 36428 MW; E7FADA0FF1BC95F9 CRC64;

Query Match 73.8%; Score 1418.5; DB 6; Length 338;
Best Local Similarity 79.5%; Pred. No. 6.4e-121;
Matches 263; Conservative 27; Mismatches 30; Indels 11; Gaps 2;

QY 1 STIIHQORIVKLPSEKELPMTQSSNGTSEPMRDRSEQHNGELRGMTINSIGR 60
Db 18 STIFHVQORLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGMTWTFINALGR 66

```

```
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLND 120
Db 67 LGNQGEYATLYALAKMNGRPAPFPAQMHSTLAPFRITLPVLHSDTASRIQWNYHLND 126
QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 127 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 186
QY 181 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 240
Db 187 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 246
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 300
Db 247 NIDTSHGDVVFAGDIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 306
QY 301 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 331
Db 307 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 337

RESULT 12
O77485 PRELIMINARY: PRT: 343 AA.
AC O77485;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL: AB015634; BAA31127.1; -.
DR InterPro: IPR002516; GI_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;
```

```
Query Match 73.8%; Score 1418.5; DB 6; Length 343;
Best Local Similarity 79.5%; Pred. No. 6.5e-121;
Matches 263; Conservative 27; Mismatches 30; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKLQPLSEKELPMTTQMSSGNTESPMRDSQHGNGELRGMTINSIGR 60
Db 23 STIFVQOQLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGWTINAIGR 71
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLND 120
Db 72 LGNQGEYATLYALAKMNGRPAPFPAQMHSTLAPFRITLPVLHSDTASRIQWNYHLND 131
QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 132 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 191
QY 181 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 240
Db 192 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 251

RESULT 14
O29505 PRELIMINARY: PRT: 347 AA.
AC O29505;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
```

```
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 300
Db 252 NIDTSHGDVVFAGDIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 311
QY 301 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 331
Db 312 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 342

RESULT 13
O9TUD3 PRELIMINARY: PRT: 343 AA.
AC O9TUD3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Alpha(1,2)fucosyl transferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALEXIS;
RX MEDLINE=20188794; PubMed=10723735;
RA Apoll P.-A., Roubinet F., Despia S., Mollicone R., Oriol R.,
RA Blancher A.;
RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
between an intronic Alu-Y element and red cell expression of ABH
antigens.";
RL Mol. Biol. Evol. 17:337-351(2000).
DR EMBL: AF080606; AAF14068.1; -.
DR InterPro: IPR002516; GI_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase.
SQ SEQUENCE 343 AA; 39001 MW; A753375D47AE8C8C CRC64;

Query Match 73.8%; Score 1417.5; DB 6; Length 343;
Best Local Similarity 79.5%; Pred. No. 8.1e-121;
Matches 263; Conservative 26; Mismatches 31; Indels 11; Gaps 2;

QY 1 STIIHQQRIVKLQPLSEKELPMTTQMSSGNTESPMRDSQHGNGELRGMTINSIGR 60
Db 23 STIFVQOQLAKIQAM--WELPV-----QIPVLASTSKALGPSQLRGWTINAIGR 71
QY 61 LGNQGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLND 120
Db 72 LGNQGEYATLYALAKMNGRPAPFPAQMHSTLAPFRITLPVLHSDTASRIQWNYHLND 131
QY 121 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQ 180
Db 132 WMEERYRHIPGHVFRFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSR 191
QY 181 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 240
Db 192 PSTFVGHVRRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPFVVTSGMAWCREE 251
QY 241 NINASRGDVVFAGNGIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 300
Db 252 NIDTSHGDVVFAGDIEGSPAKDFALLTCQNTHTIMTIGTFGIWAAYLAGGDTIYLANITL 311
QY 301 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 331
Db 312 PDSPLKFKVPEAAFLPEWVGTPADLSPLK 342

RESULT 14
O29505 PRELIMINARY: PRT: 347 AA.
AC O29505;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 10:20:58 ; Search time 2880.28 seconds
(without alignments)
10791.256 Million cell updates/sec

Title: US-10-040-863-9
Perfect score: 1068
Sequence: 1 ctccagcaggaatagtaa.....gaagagccttcgatgggaa 1068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vt.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vt.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rnd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1068	100.0	1068	10	AF042743 Rattus no
2	1068	100.0	1149	10	AF264005 Rattus no
3	1054.4	98.7	1847	10	AF131238 Rattus no
4	1052.8	98.6	2984	10	AB006138 Rattus no
5	1040.2	97.4	150840	2	AC121211 Rattus no
6	857.2	80.3	6762	10	AF214656 Mus muscu
7	857.2	80.3	206146	2	AC073774 Mus muscu
8	791.2	74.1	1044	10	AF064792 Mus muscu
9	784.6	71.6	963	10	AB039114 Mus muscu
10	784.6	71.6	963	10	AB039115 Mus muscu
11	784.6	71.6	963	10	AB039116 Mus muscu
12	784.6	71.6	963	10	AB039117 Mus muscu
13	784.6	71.6	963	10	AB039118 Mus muscu
14	784.6	71.6	963	10	AB039119 Mus muscu
15	784.6	71.6	963	10	AB039120 Mus muscu
16	784.6	71.6	963	10	AB039121 Mus muscu
17	784.6	71.6	963	10	AB039122 Mus muscu
18	761.4	71.3	963	10	AB039123 Mus spici
19	669.6	62.7	1555	10	AF131239 Rattus no
20	659.6	61.8	1107	10	AB039213 Mus spici
21	658	61.6	1107	10	AB039209 Mus muscu
22	654.8	61.3	1107	10	AB039212 Mus muscu
23	654.8	61.3	1107	10	AF113532 Mus muscu
24	654.8	61.3	1589	10	MMSEC1 Y09852 M.musculus
25	653.2	61.2	1107	10	AB039206 Mus muscu
26	653.2	61.2	1107	10	AB039207 Mus muscu
27	653.2	61.2	1107	10	AB039208 Mus muscu
28	653.2	61.2	1107	10	AB039211 Mus muscu
29	653.2	61.2	4705	10	AF214657 Mus muscu
30	651.6	61.0	1107	10	AB039210 Mus muscu
31	650	60.9	1107	10	AB039204 Mus muscu
32	650	60.9	1107	10	AB039205 Mus muscu
33	625.6	58.6	5357	10	AF214658 Mus muscu
34	614.8	57.6	2315	4	AF136895 AF136895 Sus scrofa
35	614.8	57.6	2922	4	SSU70881 AR212339 Sequence
36	612.8	57.4	1043	6	AR212339 Sequence
37	612.8	57.4	1043	6	AX029027 Sequence
38	612.8	57.4	1044	4	AF027304 Sus scrofa
39	612.2	57.3	1317	4	OCA12PT X91269 O.cuniculus
40	611.2	57.2	1043	6	AX029101 Sequence
41	594.4	55.7	1062	9	AB015635 Gorilla g
42	592.8	55.5	1062	9	AF080606 Gorilla g
43	592	55.4	1738	4	BF0742GEN X99620 Bos taurus
44	591.2	55.4	1062	9	D87934 Chlorocebus
45	591	55.3	119631	9	AC008888 Homo sapi

ALIGNMENTS

RESULT 1
AF042743 AF042743 1068 bp mRNA linear ROD 01-MAY-2001
LOCUS Rattus norvegicus alpha 1,2 fucosyltransferase mRNA, partial cds.
DEFINITION AF042743
ACCESSION AF042743
VERSION AF042743.1 GI:2801824
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1068)
AUTHORS Sherwood,A.L. and Holmes,E.H.
TITLE Cloning and expression of the catalytic domain from rat hepatoma

H35 cell GDP-fucose:GMI alpha 1-->2fucosyltransferase, an enzyme which is activated during early stages of chemical carcinogenesis in rat liver

JOURNAL Arch. Biochem. Biophys. 355 (2), 215-221 (1998)

MEDLINE 98342056

PUBMED 9675030

REFERENCE 2 (bases 1 to 1068)

AUTHORS Sherwood,A.L. and Holmes,E.H.

TITLE Direct Submission

JOURNAL Submitted (13-JAN-1998) Cell Surface Biochem., Northwest Hospital,

120 Northgate Plaza, Suite 230, Seattle, WA 98125, USA

Location/Qualifiers

1. .1068

source

/organism="Rattus norvegicus"

/strain="Buffalo rat"

/db_xref="taxon:10116"

/cell_line="hepatoma Reuber H35"

<1. .1068

/codon_start=1

/product="alpha 1,2 fucosyltransferase"

/protein_id="AAC14695.1"

/db_xref="GI:2801825"

/translation="LQRIIVKLQPLSEKELPMTTOMSSGNTSEPMRRDSQHNGEL

RGMTINISGLGNMGEYRIHPIFALRMNGRIAPISAMNHALAPIRISLPVLHSDT

AKTIQWNYHLNDWMEERYRHPGFHYFTGPGSWTFYHLRPEILKEFTLHDHRE

EAQAFRLRLVNGSQPTFVGVHVRGDYVHVPNVKGVADVRLGYLEKALDMFRARY

SSPVEVYTSNGMACRININASRGDVFAGNIEGSPAKDFALLTQCNHTIMTIGTFG

IWAAYLAGGDITLYLANTLDPSPFLKVPKPEAFLPEWVGIPADLSPLLKALITACPR

SHFLKAKGVTCYVAGRAF"

BASE COUNT 237 a 308 c 300 g 223 t

ORIGIN

Query Match 100.0%; Score 1068; DB 10; Length 1068;

Best Local Similarity 100.0%; Pred. No. 2.9e-270;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGACGACGGAATAGTAGCTCAACCCCTGTGACAGAGAGGAATACCGATGACGACT 60
Db 1 CTCGACGACGGAATAGTAGCTCAACCCCTGTGACAGAGAGGAATACCGATGACGACT 60
QY 61 CAATGTCCTCGGAAACACAGAACCCAGATGCGACGGGACGAGCAGCAGTGGG 120
Db 61 CAATGTCCTCGGAAACACAGAACCCAGATGCGACGGGACGAGCAGCAGTGGG 120
QY 121 AATGAGAGCTGCGGGGCGATGTCACGATCAATTCATTGCGCGCTGGGGAACAGATG 180
Db 121 AATGAGAGCTGCGGGGCGATGTCACGATCAATTCATTGCGCGCTGGGGAACAGATG 180
QY 181 GCGGAATACGCGACACACTTTTGCACTGCGGACGAGTGAACGGACGCGCTTGCCTTCATCCCC 240
Db 181 GCGGAATACGCGACACACTTTTGCACTGCGGACGAGTGAACGGACGCGCTTGCCTTCATCCCC 240
QY 241 GCATCCATGACACAGCTCTAGCGCCCATCTTCAGGATCAGCCCTCCCGGTGTACACAGC 300
Db 241 GCATCCATGACACAGCTCTAGCGCCCATCTTCAGGATCAGCCCTCCCGGTGTACACAGC 300
QY 301 GACAGCGGCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 360
Db 301 GACAGCGGCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 360
QY 361 TACGCCACATTCGGGACACTTTGTGCGGTTCACGGGATACCCCTGCATGACACCGTGGGGAG 480
Db 361 TACGCCACATTCGGGACACTTTGTGCGGTTCACGGGATACCCCTGCATGACACCGTGGGGAG 480
QY 481 GAGGCCAGGCGCTTCTCGTGGTCTGCGGTGAATGGAGCCAGCCAGTACTTTTGTG 540
Db 481 GAGGCCAGGCGCTTCTCGTGGTCTGCGGTGAATGGAGCCAGCCAGTACTTTTGTG 540
QY 541 GGTGTCCATGTGCGCGGAGGGGACTATGTGATGTGATGCTATGCTATGCTGGAAGGCGGTG 600

Db 541 GGTGTCCATGTGCGCGGAGGGAGCTATGTGATGCTATGCTATGCTGGAAGGCGGTG 600
QY 601 CTGGCTGACCGGGGTACCTGGAAGGCCCTGGATATGTCGGGCGACGCTATTCATCT 660
Db 601 CTGGCTGACCGGGGTACCTGGAAGGCCCTGGATATGTCGGGCGACGCTATTCATCT 660
QY 661 CCAGTCTTCGTGGTTTACAGCAAGCGTATGCTGCTGCGGGGAGCAACATTAATGCTTCC 720
Db 661 CCAGTCTTCGTGGTTTACAGCAAGCGTATGCTGCTGCGGGGAGCAACATTAATGCTTCC 720
QY 721 CGAGGAGACGTGGTGTTCGCGGCAATGATATGAGGGGTGCGCCAGCCAGGACTTCGCG 780
Db 721 CGAGGAGACGTGGTGTTCGCGGCAATGATATGAGGGGTGCGCCAGCCAGGACTTCGCG 780
QY 781 CTGCTCACCAGTGCACACCAACCATCATGACTATTTGGGACCTTTGGGCTGCC 840
Db 781 CTGCTCACCAGTGCACACCAACCATCATGACTATTTGGGACCTTTGGGCTGCC 840
QY 841 TACCTGGCAGGTGGTATACCATCTACTAGGCAACTACACCTTCCGGATTCCTCGTTC 900
Db 841 TACCTGGCAGGTGGTATACCATCTACTAGGCAACTACACCTTCCGGATTCCTCGTTC 900
QY 901 CTCAAAGTCTTTAAGCCAGAGCGCCCTTCTACCCGAATGGTGGGATCCCTCCGAT 960
Db 901 CTCAAAGTCTTTAAGCCAGAGCGCCCTTCTACCCGAATGGTGGGATCCCTCCGAT 960
QY 961 CTGTCGCCACCTTCCTTAAGGCAATACACACCGCTGCTCGGTGCCACCTCCACCTCAAG 1020
Db 961 CTGTCGCCACCTTCCTTAAGGCAATACACACCGCTGCTCGGTGCCACCTCCACCTCAAG 1020
QY 1021 GCAAAAGAGTCACTTGTACGTCGACAGAGAGCGCTTCTGATGGAA 1068
Db 1021 GCAAAAGAGTCACTTGTACGTCGACAGAGAGCGCTTCTGATGGAA 1068

RESULT 2

AF264005

LOCUS AF264005

DEFINITION Rattus norvegicus alpha 1-2 fucosyltransferase mRNA, complete cds.

ACCESSION AF264005

VERSION AF264005.1

KEYWORDS GI:8099876

SOURCE Rattus norvegicus.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 1149)

Sherwood,A.L., Stroud,M.R., Lavery,S.B. and Holmes,E.H.

An amino acid region at the N-terminus of rat hepatoma alpha1-->2

fucosyltransferase modulates enzyme activity and interaction with

lipids: strong preference for glycosphingolipids containing

terminal galactose-->3galNAc-structures

Biochemistry 40 (19), 5708-5719 (2001)

21240119

PUBMED 11341836

2 (bases 1 to 1149)

Sherwood,A.L. and Holmes,E.H.

Direct Submission

Submitted (05-MAY-2000) Molecular Medicine, Northwest Hospital,

2203 Airport Way South, Seattle, WA 98134, USA

Location/Qualifiers

1. .1149

/organism="Rattus norvegicus"

/strain="Fischer"

/db_xref="taxon:10116"

/cell_line="Reuber H35 hepatoma"

1. .1143

/note="glycosyltransferase"

/codon_start=1

/product="alpha 1-2 fucosyltransferase"

/protein_id="AAF72200.1"

CDS

```
/db_xref="GI:8099677"
/translation="MASAQVPPSFPLAHLEIFVFTSTIIHLQQRIVKLOPLSEKELP
MTTOMSSNGTESPEMRDSEQHNGELRGMEFTINSIGRLNGOMGEVATLFAIARXNGR
LAFIPASHMNAIAPIFRISLPVLHSDTAKKIPQNYHLNDWEERYRHIPGFHVRFTG
YPCSTFYVHLHLEPEILKEFTLHDHVRERQAFLRLGRVNGSQPSTFVGVHVRGDDYVH
VMPNVKGVVADRGYLEKALDMFRARYSSPVFTVSNMGAWCRENINASRGDVFAGN
GIEGSPAKDFALLTCNHTIMTIGTFGIWAAYLAGSDTIYLANITLPSDFFLKVKPE
AAFLPEWVGIPADLSPLKALPAPCRSHFHLKAKGVTCYVAGRAF"
BASE COUNT 247 a 341 c 311 g 250 t
ORIGIN
Query Match 100.0%; Score 1068; DB 10; Length 1149;
Best Local Similarity 100.0%; Pred. No. 2.9e-270;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTCTAGAGNAGGAATACCGATGACGACT 60
|||||
DB 82 CTCACGACGCAATAGTGAAGCTCCAAACCCCTGTCTAGAGNAGGAATACCGATGACGACT 141
QY 61 CAATGTCTCTCGGGAACACAGAAAGCCAGAGATCGAGCGGACACGAGGAGCATGGG 120
|||||
DB 142 CAATGTCTCTCGGGAACACAGAAAGCCAGAGATCGAGCGGACACGAGGAGCATGGG 201
QY 121 AATGAGACCTCGGGGCAATGTCAGGATCAATTCATTGGCCGGCTGGGGAACCAAGATG 180
|||||
DB 202 AATGAGAGCTCGGGGCAATGTCAGGATCAATTCATTGGCCGGCTGGGGAACCAAGATG 261
QY 181 GCGGAATACGCCACACTCTTTCACCTGGCCAGGATGAACGAGCGCTTCGGTTCATCCCC 240
|||||
DB 262 GCGGAATACGCCACACTCTTTCACCTGGCCAGGATGAACGAGCGCTTCGGTTCATCCCC 321
QY 241 GCATCATCCACAACGCTCTAGCGCCCACTTCAGGATCAGCTCCCGGTGTTACACAGC 300
|||||
DB 322 GCATCATCCACAACGCTCTAGCGCCCACTTCAGGATCAGCTCCCGGTGTTACACAGC 381
QY 301 GACAGGCCAAAGATGCCATGGCAGAAATACCATCTCAAGCTCGATGAGGAGCGT 360
|||||
DB 382 GACAGGCCAAAGATGCCATGGCAGAAATACCATCTCAAGCTCGATGAGGAGCGT 441
QY 361 TACCGGCACATTCGGGACACTTGTGGCTCTACGGGATACCGGTCCTCGGACCTTC 420
|||||
DB 442 TACCGGCACATTCGGGACACTTGTGGCTCTACGGGATACCGGTCCTCGGACCTTC 501
QY 421 TACACACCTCGGCCACAGATCTCTGAAGAGTTCACCCCTGCATGACACACGTGGGGAG 480
|||||
DB 502 TACACACCTCGGCCACAGATCTCTGAAGAGTTCACCCCTGCATGACACACGTGGGGAG 561
QY 481 GAGGCCAGGCTTCTCGGTGCTGCGGTGAATGGAGCGACGACCGAGTACTTTTGTG 540
|||||
DB 562 GAGGCCAGGCTTCTCGGTGCTGCGGTGAATGGAGCGACGACCGAGTACTTTTGTG 621
QY 541 GGTGTCCATGTGCGCGGAGGGGACTATGTGCATGTTCATCCCTAAATGTGTGAAGGCGGTG 600
|||||
DB 622 GGTGTCCATGTGCGCGGAGGGGACTATGTGCATGTTCATCCCTAAATGTGTGAAGGCGGTG 681
QY 601 GTGGCTGACCGGGGTACTCTGGAAGAGCCCTGGATATGTCGGGACCGCTATTCACT 660
|||||
DB 682 GTGGCTGACCGGGGTACTCTGGAAGAGCCCTGGATATGTCGGGACCGCTATTCACT 741
QY 661 CCAGTCTCGGTGTACAGCAACGGTATGCGCTGGTGGCGGAGACACATTAATGCTTCC 720
|||||
DB 742 CCAGTCTCGGTGTACAGCAACGGTATGCGCTGGTGGCGGAGACACATTAATGCTTCC 801
QY 721 CGAGGAGACGTGGTGTTCGCGGGCAATGTAATGAGGGGTGCGCCAGCCAAAGACTTCGG 780
|||||
DB 802 CGAGGAGACGTGGTGTTCGCGGGCAATGTAATGAGGGGTGCGCCAGCCAAAGACTTCGG 861
QY 781 CTGCTCACCCACTGCAACACACACCATCATGACTATGGACCTTTGGGATTTGGGCTGCC 840
|||||
DB 862 CTGCTCACCCACTGCAACACACACCATCATGACTATGGACCTTTGGGATTTGGGCTGCC 921
QY 841 TACCTGGGAGGTGGTATACCATCATGACTATGGACCTTTAGCCACTACACCTTCGCGATTC 900
|||||
```

```
922 TACCTGGCAGGTGGTGATACCATCTACTTAGCAACACTACACCCCTTCGGGATTCCTCGTTC 981
901 CTCAAAGTCTTTAAGCCAGAGGAGCGCTTCCTACCCGAATGGGTGGGCATCCCTGCCGAT 960
|||||
982 CTCAAAGTCTTTAAGCCAGAGGAGCGCTTCCTACCCGAATGGGTGGGCATCCCTGCCGAT 1041
|||||
961 CTGTCGCCACTCTCTAAGGATTAACACAGACCTGTCTCGTGGTCCCACTTCCACCTCAAG 1020
|||||
1042 CTGTCGCCACTCTCTAAGGATTAACACAGACCTGTCTCGTGGTCCCACTTCCACCTCAAG 1101
QY 1021 GCAAAAGGAGTCACTTGTACGTGCGAGGAGAGCGCTTCGTGATGGGAA 1068
|||||
DB 1102 GCAAAAGGAGTCACTTGTACGTGCGAGGAGAGCGCTTCGTGATGGGAA 1149
RESULT 3
AF131238 1847 bp DNA linear ROD 05-MAR-2001
LOCUS Rattus norvegicus alpha 1,2-fucosyltransferase B (FTB) gene,
DEFINITION complete cds.
ACCESSION AF131238
VERSION AF131238.1 GI:4580672
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1847)
Bureau, V., Marionneau, S., Caillean-Thomas, A., Le Moullac-Vaidye, B.,
Liehr, T. and Le Pendu, J.
Comparison of the three rat GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferases FTA, FTB and FTC
Eur J. Biochem. 268 (4), 1006-1019 (2001)
21099374
PUBMED 11179967
REFERENCE 2 (bases 1 to 1847)
Bureau, V., Le Moullac-Vaidye, B., Liehr, T., Denis, M. and Le Pendu, J.
Direct Submission
AUTHORS Submitted (25-FEB-1999) INSERM U419, Institut de Biologie, 9 Quai
Moncoussu, Nantes 44035, France
FEATURES
Location/Qualifiers
1..1847
/organism="Rattus norvegicus"
/strain="BDIX"
/db_xref="taxon:10116"
/chromosome="1"
/map="1q22-q31"
<211..>1275
/gene="FTB"
<211..>1275
/gene="FTB"
/product="alpha 1,2-fucosyltransferase B"
211..1275
/gene="FTB"
/notes="glycosyltransferase; alpha 2-fucosyltransferase B"
/codon_start=1
/product="alpha 1,2-fucosyltransferase B"
/protein_id="A4D24469.1"
/db_xref="GI:4580673"
/translation="MASAQVPPSFPLAHLEIFVFTSTIIHLQQRIVKLOPLSEKELP
MTTOMSSNGTESPEMRDSEQHNGELRGMEFTINSIGRLNGOMGEVATLFAIARXNGR
LAFIPASHMNAIAPIFRISLPVLHSDTAKKIPQNYHLNDWEERYRHIPGFHVRFTG
YPCSTFYVHLHLEPEILKEFTLHDHVRERQAFLRLGRVNGSQPSTFVGVHVRGDDYVH
VMPNVKGVVADRGYLEKALDMFRARYSSPVFTVSNMGAWCRENINASRGDVFAGN
GIEGSPAKDFALLTCNHTIMTIGTFGIWAAYLAGSDTIYLANITLPSDFFLKVKPE
AAFLPEWVGIPADLSPLKALPAPCRSHFHLKAKGVTCYVAGRAF"
BASE COUNT 452 a 503 c 457 g 435 t
ORIGIN
Query Match 98.7%; Score 1054.4; DB 10; Length 1847;
Best Local Similarity 99.8%; Pred. No. 1.1e-266;
Matches 1066; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```



```
QY 1 CTCACGACGCAATAGTGAAGCTCAACCCCTGTCAGAGAGGAATATACCGATGACGACT 60
DB |||||||
QY 292 CTCACGACGCAATAGTGAAGCTCAACCCCTGTCAGAGAGGAATATACCGATGACGACT 351
DB |||||||
QY 61 CAAATGTCCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGCAGCAGCATGGG 120
DB |||||||
QY 352 CAAATGTCCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGCAGCAGCATGGG 411
DB |||||||
QY 121 AATGAGAGCTGCGGGGATGTTACGATCAATTCATTTGCCGCTGGGAGAACAGATG 180
DB |||||||
QY 412 AATGAGAGCTGCGGGGATGTTACGATCAATTCATTTGCCGCTGGGAGAACAGATG 471
DB |||||||
QY 181 GCGGAATACGCCACACTCTTTGCACTGCCAGGATGAACGAGGCTTGGCTTCACTCCC 240
DB |||||||
QY 472 GCGGAATACGCCACACTCTTTGCACTGCCAGGATGAACGAGGCTTGGCTTCACTCCC 531
DB |||||||
QY 241 GCATCCATGCACAAGCTCTAGCGGCCATCTTCAGGATCAGCCTCCCGTGTACACAGC 300
DB |||||||
QY 532 GCATCCATGCACAAGCTCTAGCGGCCATCTTCAGGATCAGCCTCCCGTGTACACAGC 591
DB |||||||
QY 301 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAGAGCTGATGAGAGCGT 360
DB |||||||
QY 592 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAGAGCTGATGAGAGCGT 651
DB |||||||
QY 361 TACCGCCACACTTCGCGGACACTTTGCGGCTTCACGGGATACCCGCTGCTGGACCTTC 420
DB |||||||
QY 652 TACCGCCACACTTCGCGGACACTTTGCGGCTTCACGGGATACCCGCTGCTGGACCTTC 711
DB |||||||
QY 421 TACCAACCACTGCGGCCAGAGATCCTGAAGAGTTACCCCTGCATGACCACTGCGGGAG 480
DB |||||||
QY 712 TACCAACCACTGCGGCCAGAGATCCTGAAGAGTTACCCCTGCATGACCACTGCGGGAG 771
DB |||||||
QY 481 GAGCGCCAGGCTTCCTGCGTGGCTGCGGCTGATGGAGCCAGCCAGTACTTTTGTG 540
DB |||||||
QY 772 GAGCGCCAGGCTTCCTGCGTGGCTGCGGCTGATGGAGCCAGCCAGTACTTTTGTG 831
DB |||||||
QY 541 GGTGTCTCATGTGCGCGGAGGACTATGTGATGTCATGCTCAATGTGTGAAGGCGGTG 600
DB |||||||
QY 832 GGTGTCTCATGTGCGCGGAGGACTATGTGATGTCATGCTCAATGTGTGAAGGCGGTG 891
DB |||||||
QY 601 GTGGCTGACCGGGTTACTCTGAAAGCCCTGGATATGTTCCGGGACAGTATTCATCT 660
DB |||||||
QY 892 GTGGCTGACCGGGTTACTCTGAAAGCCCTGGATATGTTCCGGGACAGTATTCATCT 951
DB |||||||
QY 661 CCAGTCTTCGCGTTTACAAGCAAGGATGCGCTGTCGCGGAGAAACATTAATGCTTCC 720
DB |||||||
QY 952 CCAGTCTTCGCGTTTACAAGCAAGGATGCGCTGTCGCGGAGAAACATTAATGCTTCC 1011
DB |||||||
QY 721 CGAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTGCGGCAAGGACTTCCGG 780
DB |||||||
QY 1012 CGAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTGCGGCAAGGACTTCCGG 1071
DB |||||||
QY 781 CTGCTCACCCAGTGCACACCACTATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 840
DB |||||||
QY 1072 CTGCTCACCCAGTGCACACCACTATGACTATTTGGGACCTTTGGGATTTGGGCTGCC 1131
DB |||||||
QY 841 TACCTGGCAGCTGGTGATACCATCTACTTACCCACTACACCTTCGCGGATTCCTCGGTC 900
DB |||||||
QY 1132 TACCTGGCAGCTGGTGATACCATCTACTTACCCACTACACCTTCGCGGATTCCTCGGTC 1191
DB |||||||
QY 901 CTCGAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGGTGGGATCCCTTCCCGAT 960
DB |||||||
QY 1192 CTCGAAGTCTTTAAGCCAGAGCAGCCTTCCTACCCGAATGGGTGGGATCCCTTCCCGAT 1251
DB |||||||
QY 961 CTGTCCTCCACTGCTTAAGGCAATTAACACCACTGCTGCTCGGTCCTTCCACTTCAAG 1020
DB |||||||
QY 1252 CTGTCCTCCACTGCTTAAGGCAATTAACACCACTGCTGCTCGGTCCTTCCACTTCAAG 1310
DB |||||||
QY 1021 GCAAAAGAGTCACTTTGCTAGCTGCGCAGAGAGGCTTCTGATGGAA 1068
DB |||||||
QY 1311 GCAAAAGAGTCACTTTGCTAGCTGCGCAGAGAGGCTTCTGATGGAA 1358
DB |||||||
```

```
RESULT 4
LOCUS AB006138
DEFINITION Rattus norvegicus FTB mRNA for alpha 1,2-fucosyltransferase, complete cds.
ACCESSION AB006138
VERSION AB006138.1 GI:2317265
KEYWORDS FTB; alpha 1,2-fucosyltransferase.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (sites)
AUTHORS Soejima,M., Wang,B., Koda,Y. and Kimura,H.
TITLE Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-fucosyltransferase genes
JOURNAL Unpublished
AUTHORS Koda,Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-1997) Yoshiro Koda, Kurume University, School of Medicine, Department of Forensic Medicine, Asahimachi 57, Kurume, Fukuoka 830, Japan (E-mail:y.koda@med.kurume-u.ac.jp, Tel:0942-31-7554, Fax:0942-31-7700)
FEATURES
source location/Qualifiers
1..2984
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="1"
/map="1q22-q31"
/cell_line="RCN-9"
/tissue_type="colon cancer"
1..2984
/gene="FTB"
214..1278
/gene="FTB"
/codon_start=1
/product="alpha 1,2-fucosyltransferase"
/protein_id="BAA21742.1"
/db_xref="GI:2317265"
/translation="MASAQVFPFSPFLAHFLIFVFTSTIIHQORIVKLOPLSKELP
MTTOMSSGNTSPEMRRDSEOHGELGRMTINSIGLNGMGEYATLFLALRNMR
LAFIPASMHNALAPIERISLPVLHSDTAKKIPWONYHLNDWERYRIPIGHFVFTG
YPCSWTFYHLKRLKEFTLHDVREDAQFLRLGLRVNGSQPTFVGVHVRGDIYH
VPMNVKGVADRGYLEKALDPRARYSPVFTVTSNGMWCNRINASRGDVFAGN
GIEGSPKDFALLTQCNTINTIGTGTGIAAAYLAGGDTIYLANVTLPDSPFLKFKPE
AAFLPEWVGIPADLSPLKH"
BASE COUNT 729 a 808 c 761 g 586 t
ORIGIN
Query Match 98.6%; Score 1052.8; DB 10; Length 2984;
Best Local Similarity 99.7%; Pred. No. 2.9e-266;
Matches 1065; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 CTCACGACGCAATAGTGAAGCTCAACCCCTGTCAGAGAGGAATATACCGATGACGACT 60
DB 295 CTCACGACGCAATAGTGAAGCTCAACCCCTGTCAGAGAGGAATATACCGATGACGACT 354
QY 61 CAAATGTCCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGCAGCAGCATGGG 120
DB 355 CAAATGTCCTCGGGAACACAGAAAGCCAGAGATGCCAGCGGACAGCAGCAGCATGGG 414
QY 121 AATGAGAGCTGCGGGGATGTTACGATCAATTCATTTGCCGCTGGGAGAACAGATG 180
DB 415 AATGAGAGCTGCGGGGATGTTACGATCAATTCATTTGCCGCTGGGAGAACAGATG 474
QY 181 GCGGAATACGCCACACTCTTTGCACTGCCAGGATGAACGAGGCTTGGCTTCACTCCC 240
DB 475 GCGGAATACGCCACACTCTTTGCACTGCCAGGATGAACGAGGCTTGGCTTCACTCCC 534
QY 241 GCATCCATGCACAAGCTCTAGCGGCCATCTTCAGGATCAGCCTCCCGTGTACACAGC 300
```

Db	535	GCATCCATGCACACGCTCTAGCGCCCATCTTCAGGATACGCTCCCGGTITACACAGC	594	
Qy	301	GACAGGCCAAAAGATCCCATGGCAGAAATACATCTCAACGACTGGATGGAGGACGCT	360	
Db	595	GACAGGCCAAAAGATCCCATGGCAGAAATACATCTCAACGACTGGATGGAGGACGCT	654	
Qy	361	TACCGCCACATTCGCGGACACTTTGTCGCTGTACGGGATACCGTGCTCTCTGACACCTTC	420	
Db	655	TACCGCCACATTCGCGGACACTTTGTCGCTGTACGGGATACCGTGCTCTCTGACACCTTC	714	
Qy	421	TACACACACCTGCGCCGACAGATCTCTGAAGAGATTCACCTGATGACACCGTCCGGAG	480	
Db	715	TACACACACCTGCGCCGACAGATCTCTGAAGAGATTCACCTGATGACACCGTCCGGAG	774	
Qy	481	GAGGCCACGCGCTTCTGCGGTGCTCGGGTGAATGGGAGCCAGGACTACTTTTGTG	540	
Db	775	GAGGCCACGCGCTTCTGCGGTGCTCGGGTGAATGGGAGCCAGGACTACTTTTGTG	834	
Qy	541	GGTGCTCATGTGCGCGAGGGACTATGTGCATGTGATGCTTATGTGTGGAAGGGCGTG	600	
Db	835	GGTGCTCATGTGCGCGAGGGACTATGTGCATGTGATGCTTATGTGTGGAAGGGCGTG	894	
Qy	601	GTGCTGACCGGGGTTACCTGGAAGGCCCTGGATATGTTCCGGGACGCTATTCATCT	660	
Db	895	GTGCTGACCGGGGTTACCTGGAAGGCCCTGGATATGTTCCGGGACGCTATTCATCT	954	
Qy	561	CCATGCTCTGTTTACACGCAACGGTATGCTTGTGCGGGAGACATTAATGCTTCC	720	
Db	955	CCATGCTCTGTTTACACGCAACGGTATGCTTGTGCGGGAGACATTAATGCTTCC	1014	
Qy	721	CGAGGACGCTGTTGTCGCGGCAATGGTATTCAGGGGTGCCAGCAAGGACTTCGCG	780	
Db	1015	CGAGGACGCTGTTGTCGCGGCAATGGTATTCAGGGGTGCCAGCAAGGACTTCGCG	1074	
Qy	781	CTGCTCACCCAGTGCACACCAACCATCATGACTATTTGGGACTTTGGGATTTGGGCTGCC	840	
Db	1075	CTGCTCACCCAGTGCACACCAACCATCATGACTATTTGGGACTTTGGGATTTGGGCTGCC	1134	
Qy	841	TACCTGGCAGTGGTGATACCATCTACTTACCACTACACCTTCCGGATTCCTCGTTC	900	
Db	1135	TACCTGGCAGTGGTGATACCATCTACTTACCACTACACCTTCCGGATTCCTCGTTC	1194	
Qy	901	CTCAAGTCTTTAAGCCAGAGCGACCTTCTTACCCGAATGGTGGGATCCCTCCCGAT	960	
Db	1195	CTCAAGTCTTTAAGCCAGAGCGACCTTCTTACCCGAATGGTGGGATCCCTCCCGAT	1254	
Qy	961	CTGTCGCCACTTCCTTAAGGCATTAACACGACCTGCTCCGTCGTCCTTCCACCTCAAG	1020	
Db	1255	CTGTCGCCACTTCCTTAAGGCATTAACACGACCTGCTCCGTCGTCCTTCCACCTCAAG	1313	
Qy	1021	GCAAAAGAGTCACTTGTAGTGTCCAGAGAGGCTTCTGATGGAA	1068	
Db	1314	GCAAAAGAGTCACTTGTAGTGTCCAGAGAGGCTTCTGATGGAA	1361	
RESULT	5			
AC121211				
LOCUS		150840 bp	DNA	linear
DEFINITION				HTG 24-JUL-2002
				Rattus norvegicus clone CH230-309A14, *** SEQUENCING IN PROGRESS
				***, 63 unordered pieces.
ACCESSION				AC121211
VERSION				AC121211.2
KEYWORDS				HTG; HTGS_PHASE1.
SOURCE				Rattus norvegicus.
ORGANISM				Rattus norvegicus
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
				Rattus.
REFERENCE				1 (bases 1 to 150840)
AUTHORS				Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,
				Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, T.,
				Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,

Bouck, J.,	Bowie, S.,	Brieva, M.,	Brown, E.,	Brown, M.,	Bryant, N.P.,
Buhay, C.,	Burch, P.,	Burkett, C.,	Burrell, K.L.,	Byrd, N.C.,	
Carroll, T.,	Carter, M.,	Cavazos, S.R.,	Chacko, J.,	Chavez, D.,	
Chen, G.,	Chen, R.,	Chen, Z.,	Chowdhry, I.,	Christopoulos, C.,	
Cleveland, C.D.,	Cox, C.,	Coyle, M.D.,	Dathorne, S.R.,	David, R.,	
Davila, M.L.,	Davis, C.,	Davy-Carroll, L.,	Dederich, P.A.,		
Delaney, K.R.,	Delgado, O.,	Denn, A.L.,	Ding, Y.,	Dinh, H.H.,	
Douthwaite, K.J.,	Draper, H.,	Dugan-Rocha, S.,	Durbin, K.J.,		
Earnhart, C.,	Edgar, D.,	Edwards, C.C.,	Elhaj, C.,	Escotto, M.,	
Falls, T.,	Ferraguto, D.,	Flagg, N.,	Ford, J.,	Foster, P.,	Frantz, P.,
Gabis, A.,	Geo, J.,	Garcia, A.,	Garnier, T.,	Garza, N.,	Gill, R.R.,
Gorrell, J.H.,	Guevara, W.,	Gunaratne, P.,	Hale, S.,	Hamilton, K.,	
Harris, C.,	Harris, K.,	Hart, M.,	Havlak, P.,	Hawes, A.,	Hernandez, J.,
Hernandez, O.,	Hodgson, A.,	Hogues, M.,	Holloway, C.,	Hollins, B.,	
Homs, F.,	Howard, S.,	Huber, J.,	Hulyk, S.,	Hume, J.,	Jackson, L.E.,
Jacobson, B.,	Jia, Y.,	Johnson, R.,	Jolivet, S.,	Joudah, S.,	
Karlsson, E.,	Kelly, S.,	Khan, U.,	King, L.,	Korvah, J.,	Kovar, C.,
Kratovic, J.,	Kureshi, A.,	Landry, N.,	Leal, B.,	Lewis, L.C.,	Lewis, L.,
Li, J.,	Li, Z.,	Lichtarge, O.,	Lieu, C.,	Liu, J.,	Liu, W.,
Lozano, R.J.,	Lu, X.,	Lucier, A.,	Lucier, R.,	Luna, R.,	Ma, J.,
Mareshwari, M.,	Mapua, P.,	Martin, R.,	Martindale, A.,	Martinez, E.,	
Massey, E.,	Mawhiney, E.,	McLeod, M.P.,	Meador, M.,	Mei, G.,	Metzker, M.,
Miner, G.,	Miner, Z.,	Mitchell, T.,	Mohabbat, K.,	Morgan, M.,	Morris, S.,
Moser, M.,	Neal, D.,	Newton, J.,	Newton, N.,	Nguyen, A.,	Nguyen, N.,
Nguyen, N.,	Nickerson, E.,	Nwokenkwo, S.,	Oguh, M.,	Okunolu, G.,	
Oragunye, N.,	Oviedo, R.,	Pace, A.,	Payton, B.,	Peery, J.,	Perez, L.,
Peters, L.,	Pickens, R.,	Primus, E.,	Pu, L.L.,	Quiles, M.,	Ren, Y.,
Rivera, M.,	Rojas, A.,	Rojubokan, I.,	Rolfe, M.,	Ruiz, S.,	Savery, G.,
Scherer, S.,	Scott, G.,	Shen, H.,	Shooshtari, N.,	Sisson, I.,	
Sodergren, E.,	Sonaike, T.,	Sparks, A.,	Stanley, H.,	Stone, H.,	
Sutton, J.,	Swatek, A.,	Tabor, P.,	Tamerisa, A.,	Tamerisa, K.,	Tang, H.,
Tatton, J.,	Taylor, C.,	Taylor, T.,	Telford, B.,	Thomas, N.,	Thomas, S.,
Usmani, K.,	Vasquez, L.,	Vera, V.,	Villalón, D.,	Vinson, R.,	Wang, Q.,
Wang, S.,	Ward-Moore, S.,	Warren, R.,	Washington, C.,	Watlington, S.,	
Williams, C.,	Williamson, A.,	Wleczyk, R.,	Wooden, S.,	Worley, K.,	
Wu, C.,	Wu, Y.,	Wu, Y.F.,	Zhou, J.,	Zorrilla, S.,	Nelson, D.,
Weinstock, G.	and Gibbs, R.				
Unpublished					
2 (bases 1 to 150840)					
Worley, K.C.					
Direct Submission					
Submitted (16-MAY-2002)					
of Molecular and Human Genetics, Baylor College of Medicine, One					
Baylor Plaza, Houston, TX 77030, USA					
3 (bases 1 to 150840)					
Worley, K.C.					
Direct Submission					
Submitted (24-JUL-2002)					
of Molecular and Human Genetics, Baylor College of Medicine, One					
Baylor Plaza, Houston, TX 77030, USA					
On Jul 19, 2002 this sequence version replaced gi:20806241.					
----- Genome Center					
Center: Baylor College of Medicine					
Center code: BCM					
Web site: http://www.hgsc.bcm.tmc.edu/					
Contact: hgsc-help@bcm.tmc.edu					
----- Project Information					
Center project name: GYKX					
Center clone name: CH230-309A14					
----- Summary Statistics					
Sequencing vector: Plasmid					
Chemistry: Dye-terminator Big Dye; 100% of reads					
Assembly program: Phrap; version 0.990329					
Consensus quality: 106837 bases at least Q40					
Consensus quality: 113151 bases at least Q30					
Consensus quality: 116417 bases at least Q20					

* NOTE: Estimated insert size may differ from sequence length					
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).					
* NOTE: This is a 'working draft' sequence. It currently					
* consists of 63 contigs. The true order of the pieces					
* is not known and their order in this sequence record is					


```
Db 143184 CTACCCACCTCGGCCAGAGATCTGAAGGAGTTACCCCTCATGACACCTGCGGGA 143243
QY 480 GGAGGCCAGGCTTCTCGTGGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTTGT 539
Db 143244 GGAGGCCAGGCTTCTCGTGGTCTGCGGGTGAATGGAGCCAGCCGAGTACTTTGT 143303
QY 540 GGGTGTCCATGTGCGCGGAGGGGACATATGTGCATGTGTCATGCCCTAAATGTGGAAGGCGT 599
Db 143304 GGGTGTCCATGTGCGCGGAGGGGACATATGTGCATGTGTCATGCCCTAAATGTGGAAGGCGT 143363
QY 600 GTGGGTGACCGGGGTTACTGGAAGGCCCTGGATATGTTCCGGCCAGCTATTTCATC 659
Db 143364 GTGGGTGACCGGGGTTACTGGAAGGCCCTGGATATGTTCCGGCCAGCTATTTCATC 143423
QY 660 TCCAGTCTTCTGGTGTACAGCAACCGTATGSCCTGGTCCGGGAGAACATTAATGCTTC 719
Db 143424 TCCAGTCTTCTGGTGTACAGCAACCGTATGSCCTGGTCCGGGAGAACATTAATGCTTC 143483
QY 720 CCGAGGAGACGTGGTCTTCGCGGGCAATGGTATGAGGGGTGCCAGCCAAAGACTTCGC 779
Db 143484 CCGAGGAGACGTGGTCTTCGCGGGCAATGGTATGAGGGGTGCCAGCCAAAGACTTCGC 143543
QY 780 GTGTCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGC 839
Db 143544 GTGTCTCACCCAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGC 143603
QY 840 CTACCTGGCAGTGGTGATPACCATCTACTTAGCCAACTACACCCCTCCGGATTCCTCGTT 899
Db 143604 CTACCTGGCAGTGGTGATPACCATCTACTTAGCCAACTACACCCCTCCGGATTCCTCGTT 143663
QY 900 CTTCAAGTCTTTAGCCAGAGGAGCCCTTCTACCCGAATGGTGGGATCCCTGCCGA 959
Db 143664 CTTCAAGTCTTTAGCCAGAGGAGCCCTTCTACCCGAATGGTGGGATCCCTGCCGA 143723
QY 960 TCTGTCCCACTCCCTTAAGGCATTAAACACGCTGTCTCGGTCCCACTTCCACCTCAA 1019
Db 143724 TCTGTCCCACTCCCTTAAGGCATTAAACACGCTGTCTCGGTCCCACTTCCACCTCAA 143782
QY 1020 GCAAAAGAGTCACTTGTAGTGCAGGAGAGCCCTTCTGATGGGA 1068
Db 143783 GCAAAAGAGTCACTTGTAGTGCAGGAGAGCCCTTCTGATGGGA 143831

RESULT 5
AF214656
LOCUS
DEFINITION Mus musculus alpha(1,2)fucosyltransferase FUT2 (Fut2) gene, complete cds.
ACCESSION AF214656
VERSION AF214656.1
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 6762)
AUTHORS Domino,S.B., Zhang,L. and Lowe,J.B.
TITLE Molecular cloning, genomic mapping, and expression of two secretor blood group alpha (1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract
J. Biol. Chem. 276 (26), 23748-23756 (2001)
JOURNAL 21316345
MEDLINE 11323419
PUBMED
REFERENCE 2 (bases 1 to 6762)
AUTHORS Domino,S.B. and Lowe,J.B.
TITLE Mus musculus alpha(1,2)fucosyltransferase FUT2 (Sec2) genomic sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 6762)
AUTHORS Domino,S.B. and Lowe,J.B.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1999) HHMI, University of Michigan, 1150 W. Med.
```

```
FEATURES
SOURCE
    Ctr. Dr., Ann Arbor, MI 48109-0650, USA
    1. 6762
    /organism="Mus musculus"
    /strain="129/Ola"
    /db_xref="taxon:10090"
    /chromosome="7"
    /map="23.2; between Klk1 and Ldh3"
    /cell_line="E14"
    /cell_type="ES"
    <2210..>3253
    /gene="Fut2"
    /note="Sec2"
    <2210..>3253
    /gene="Fut2"
    /product="alpha(1,2)fucosyltransferase FUT2"
    2210..3253
    /gene="Fut2"
    /EC_number="2.4.1.69"
    /note="fucosyltransferase; similar to the Mus musculus strain ICR Sec2 sequence deposited at GenBank Accession Number AF064792; similar to the Homo sapiens secretor blood group alpha(1,2)fucosyltransferase FUT2 sequences deposited at GenBank Accession Numbers D87942 and U17894"
    /codon_start=1
    /product="alpha(1,2)fucosyltransferase FUT2"
    /protein_id="AAF45146.1"
    /db_xref="GI:7288505"
    /translation="MASAQPFPSFLAHLFIHFVSTIIHLQQRIVKLQTLSEKLQ
    AVOMSPNARDTMOQSAKLGITFINSIGRLGNOMGEYALIFALARMGRFLAFIPES
    MHNLAPEIPRISLPVLHSDTARRIPWYHLNDWMEERYRHIPGOVYRFTGYPCKSTF
    YHLRPEILKEFTLHDHVRERAAQLRGLRVNGSOPSTFVGVHVRGDIVVHVPKVK
    GVVADRGYLEKALDRFRARYSPFVFTSNGMAWRENINTSLIGDVVFAGNIGESPA
    KFPALITQCNIHMTIGTFGIWAAYLAGGDTIYLANITLPDSFLKIFKPAAAFLEW
    MGIPADLSPLLLKH"
    polyA_signal 4943..4948
    polyA_signal 6333..6338
    BASE COUNT 1738 a 1682 c 1647 g 1691 t 4 others
    ORIGIN
    Query Match 80.3%; Score 857.2; DB 10; Length 6762;
    Best Local Similarity 89.3%; Pred. No. 9.2e-215;
    Matches 954; Conservative 0; Mismatches 93; Indels 21; Gaps 2;
    QY 1 CTCGAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAGAGAAGAAATACCGATGAGGACT 60
    Db 2291 CTCGAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAGAGAAGAAATAC---AGCGGTT 2347
    QY 61 CAAATGTCTCTCGGGAACACAGAAAGCCCGACGAGATGCGACGGGACAGCAGCATGGG 120
    Db 2348 CAAATGTCTCTACCAACACGCGCAGACAGACATGCGACGA-----G 2389
    QY 121 AATGAGAGCTCGGGGCAATGTTACAGATCAATTCATTTGGCCCGCTGGGGAACAGATG 180
    Db 2390 AATGCAAGCTCGAGGCAATATTCAGATCAATTCATTCGCGCCGCTGGGGAACAGATG 2449
    QY 181 GCGAATAGCCACACTCTTTGCACTGGCCAGGATGACGACGCGCTTGCCTTCATCCOC 240
    Db 2450 GCGAATAGCTACATGTTTTCACCTGGCCAGGATGAACGGTTCGGCTTCCTTCATCCT 2509
    QY 241 GCATCCATGCACACGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTTACACAGC 300
    Db 2510 GAATCCATGCACAGGCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTTACACAGC 2569
    QY 301 GACAGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGT 360
    Db 2570 GACACAGCCAGAAGATCCCGTGGCAGAAATACCACTCAACGACTGGATGGAGAGCGT 2629
    QY 361 TACGCCACATTCGGGACACTTTGTGGCTTCAGGCTTCAGGATACCGGTGCTCTCGACCTTC 420
    Db 2630 TACGCCACATTCGGGACACTTTGTGGCTTCAGGCTTCAGGATACCGGTGCTCTCGACCTTC 2689
    QY 421 TACCACCACCTCGCGCCAGAGATCCTGAAGGAGTTCAACCTGCTGATGACACAGCTGCGGAG 480
```


Db	13777	GGCGAATATGCTACATTTGTTGGACATGGCCAGGATGAACGGTCGGCTTGGCTTTCATCCCT	13711
Qy	241	GCATCATGCACAACGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTACACAGC	300
Db	13717	GAATCCATGCACAACGCTCTAGCGCCCATCTTCAGGATCAGTCTCCCGGTGTACACAGC	13658
Qy	301	GACAGGCCAAAAGATCCCATGGCAGAAATACCATCTCAAGCACTGGATGGAGAGCGT	360
Db	13657	GACACAGCCAGAGGATCCGCTGGCAGAAATACCACCTCAACGACTGGATGGAGAGCGT	13598
Qy	361	TACCGCCCAATCCGGGACATTTGTGGCTTCAGGGATACCCGTGCTCTGACCTTC	420
Db	13597	TACCGCCCAATCCGGGGAGATGTGGCTTCAGGGATACCCGTGCTCTGACCTTC	13538
Qy	421	TACCACACTGCCCCCAGAGATCCTGAAGAGTTTCACCTGCATGACCACTGCGGGAG	480
Db	13537	TACCACACTGGCCCCAGATCTGAGAGTTTACCCTGCACGACCATGTGGGTGAG	13478
Qy	481	GAGGCCAGGCTTCCTGGCTGGTCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG	540
Db	13477	GAGGCCAGGCTTTCCTGGCTGGCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG	13418
Qy	541	GGTGTCCATGTGGCGGAGGGGACATGTGCATGTTCATGTCATGTGTGGAAGGGGTG	600
Db	13417	GGTGTCCATGTGGCGGAGGGGACATGTGCATGTTCATGTCATGTGTGGAAGGGGTG	13358
Qy	601	GTGGCTGACCGGGTTTACCTGAAAAGGCCCTGGATATGTTCCGGGACAGCTATTCACT	660
Db	13357	GTGGCTGACCGGGTTTACCTGAAAAGGCCCTGGAGCTTCCGGGACAGCTATTCACT	13298
Qy	661	CCAGTCTCGTGTATACAGCAACGGTATGGCTTGGTCCGGGAGAACATTAATGCTTCC	720
Db	13297	CCAGTCTCGTGTATACAGCAACGGTATGGCTTGGTCCGGGAGAACATTAATGCTTCC	13238
Qy	721	CGAGGAGCTGTGTGTTCCGGGCAATGATTTAGGGTCCCGAGCAAGCACTTCGG	780
Db	13237	CTAGGAGAGCTGTGTGTTCCGGGCAATGATTTAGGGTCCCGAGCAAGCACTTCGG	13178
Qy	781	CTGCTCACCAGTGCACACACACCATCATGACTATTGGACCTTTGGATTTGGGCTGCC	840
Db	13177	CTGCTCACCAGTGCACACACACCATCATGACTATTGGACCTTTGGATTTGGGCTGCC	13118
Qy	841	TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCGGTTC	900
Db	13117	TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCGGTTC	13058
Qy	901	CTCAAAAGTCTTTAAGCCAGAGCAGCTTCCTACCCGATGGTGGGATCCCTGCCGAT	960
Db	13057	CTCAAAATCTTTAAGCCAGAGCAGCTTCCTCCCGAGTGGATGGCATCCCGCAGAC	12998
Qy	961	CTGTCCCCACTCTTAAGSCATTAACACAGAGCTGTGCTCGTCCCACTTCCACTCAAG	1020
Db	12997	CTGTCCCCACTCTTAAGSCATTAACACAGAGCTGTGCTCGTCCCACTTCCACTCAAG	12938
Qy	1021	GCAAAGAGTCACTTGTACCTCGCAGAGAGCGCTTCTGATGGGAA	1068
Db	12937	GCAGAAGAGTCACTTGTACATCCAGAGAGGCGCTTCTACTGGAA	12890
RESULT 8			
AF064792			
LOCUS	AF064792	1044 bp	mRNA linear ROD 24-OCT-2000
DEFINITION	Mus musculus GDP-L-fucose:beta-D-galactoside		
ACCESSION	AF064792		
VERSION	AF064792.1		
KEYWORDS	GI:3142705		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
	1. (bases 1 to 1044)		
	Lin.B., Hayashi,Y., Saito,M., Sakakibara,Y., Yanagisawa,M. and		

TITLE
 GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II, and not MFUT-I or -III, is induced in a restricted region of the digestive tract of germ-free mice by host-microbe interactions and cycloheximide
JOURNAL Biochim. Biophys. Acta 1487 (2-3), 275-285 (2000)
MEDLINE 20471982
PUBMED 11018479
REFERENCE 2 (bases 1 to 1044)
AUTHORS Lin, B., Hayashi, Y., Saito, M., Sakakibara, Y., Yanagisawa, M. and Iwamori, M.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1998) Chemistry, Faculty of Science & Technology
FEATURES Kinki University, Kowakae, Higashi-osaka 577-8502, Japan
source Location/Qualifiers
 1. .1044
 /organism="Mus musculus"
 /strain="ICR"
 /db_xref="taxon:10090"
 /tissue_type="gastrointestinal tract"
 1. .1044
 /EC_number="2.4.1.69"
 /note="sec2; Se"
 /codon_start=1
 /product="GDP-L-fucose:beta-D-galactoside
 2-alpha-L-fucosyltransferase"
 /protein_id="AAC16887.1"
 /db_xref="GI:3142706"
 /translation="MASAQVFPSPLAHLFVFTSIHLLQOOIVKLOLSEKEL
 AVQSSPNARTMQQAKGIPTINSIGRLNOMGEYATLFLARNVGRALFPE
 MHNALPFRISLVBHSDTARRIPQWYHLNDWMEERYRIPGQVRYTGYPCSWI
 YHHLRPIELKEFTLBDYVREEAQAFRLGLRVNGSQPSTFVGVHVRGDYVHVMPKVI
 GVADRYGLKALDRFRARYSPFVVTSGMAWCNRENTSLGDFVTFAGNIGESPI
 KDFALLTCNHTINTIGTFGIWAAVLAGGDIILANYTLDPSPFLKIFKPRAAFLPE
 MGIPADLSPLKH"
BASE COUNT 229 a 318 c 271 g 226 t
ORIGIN
 Query Match 74.1%; Score 791.2; DB 10; Length 1044;
 Best Local Similarity 89.9%; Pred. No. 2,2e-197;
 Matches 879; Conservative 0; Mismatches 78; Indels 21; Gaps 2
 Qy 1 CTCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTCAGAGAAGGAATTACCGATGACGACT 60
 Db 82 CTCAGCAGCAAAATAGTGAAGCTCCAAACCCCTGTCAGAGAAGGAATTACAGG---CGGT 138
 Qy 61 CAATGTCCTCGGAACACAGAAAGCCAGANGATCGAGGGACAGCGAGCAGCATGGG 120
 Db 139 CAATGTCCTCACCAACGCGGCAAGACAGACATGCAGCA-----G 180
 Qy 121 AATGGAGAGCTCGGGGCGATGTTACAGATCAATTGCCATTGGCCGGCTGGGGGAACAGATG 180
 Db 181 AGTGCCAACTGCGAGGCGNATTTACGATCAATTCATCGCCGCGCTGGGNACAGATG 240
 Qy 181 GCGGAATACGCCACACTCTTTGCACTGGCCAGGATGAACGACGCGCTTGCGTTTCATCCCC 240
 Db 241 GCGGAATATGTACATTTGTTGCACTGSCCAGGATGAACGGTGCSCCTTGCCCTTCATCCT 300
 Qy 241 GCATCCATGCACACGCTCTAGCGCCNCTCTTCAGGATCAGGCTCCCGGTTGTACACAGC 300
 Db 301 GATATCCATGCACACGCTCTAGCGCCNCTCTTCAGGATCAGGTCAGTCTCCCGGTGTACACAGC 360
 Qy 301 GACACGGCCAAAAGATCCCATGGCAGAATTACCATCTCAACGACTGGATGGAGGAGCGT 360
 Db 361 GACACGCCAGAGGATCCCGTGGCAGATTAACCACTCAACGACTGGATGGAGGAGCGT 420
 Qy 361 TACCGGCACATTCGGGACACTTTTGGCGTTTCACGGGATACCCGTTGCTTGTGGACCTTC 420
 Db 421 TACCGGCACATCCGGGCGCAGTATGTCGTTTTCACGGGATACCCGTTGCTTGTGGACCTTC 480
 Qy 421 TACACCACTTGGCCCGCAGAGATCCTGAAGGAGTTACCCCTGCATGACCAAGTGGGGGAG 480
 Db 481 TACACCACTTGGCCCGCAGAGATCCTGAAGGAGTTACCCCTGCATGACCAAGTGGGGGAG 540

Db 739 CTAGGAGACGTGGTGTTCGGGCAATGGTATTGAGGCTCACCAGCGCAAGGACTTCGCG 798
 QY 781 CTGCTCACCAGTGCACACACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 840
 Db 799 CTCCTCACCAGTGCACACACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 858
 QY 841 TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTTCCTCGTTC 900
 Db 859 TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTTCCTCGTTC 918
 QY 901 CTCAAAGTCTTTAAGCCAGAGGACGCTTCCTACCCGGAATGGGTG 945
 Db 919 CTCAAAATCTTTAAGCCAGAGGACGCTTCCTCCCGAGTGGATG 963

RESULT 10
 AB039115
 LOCUS
 DEFINITION
 Mus musculus Fut2 gene for GDP-L-fucose:beta-D-galactoside
 2-alpha-L-fucosyltransferase, partial cds, strain: BFM/2Msf.
 AB039115
 ACCESSION
 AB039115.1 GI:15822975
 VERSION
 KEYWORDS
 Fut2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
 Mus musculus (sub_species:breviostriis, strain:BFM/2Msf) DNA.
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1
 Li, Y., Kitano, T., Koide, T., Shiroyoshi, T., Moriaki, K. and
 Saitou, N.
 Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
 of Five Mus musculus subspecies
 Unpublished
 2 (bases 1 to 963)
 Li, Y. and Saitou, N.
 Direct Submission
 Submitted (29-FEB-2000) Naruya Saitou, National Institute of
 Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,
 Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
 URL: http://sayer.lab.nig.ac.jp/, Tel: 81-559-81-6790,
 Fax: 81-559-81-6789)
 Location/Qualifiers
 1..963
 /organism="Mus musculus"
 /strain="BFM/2Msf"
 /sub_species="breviostriis"
 /db_xref="taxon:10090"
 /note="sequence used for primer design based on
 AccAF064792"
 1..963
 /gene="Fut2"
 <1..>963
 /gene="Fut2"
 /EC_number="2.4.1.69"
 /note="Sec2: Se"
 /codon_start=1
 /product="GDP-L-fucose:beta-D-galactoside
 2-alpha-L-fucosyltransferase"
 /protein_id="BAB68639.1"
 /db_xref="GI:15822975"
 /translation="FLIFVFTIIHLOQRIVKLTLSKELQAVOMSPNARTDM
 QOSAKLOGIFTINSIGLQMGVATLALARNGLARIPESMHNALAFIRISLP
 VLHSDTARRIPQYHNDWMEYRIIPQYVFTGTPCSVTFYHLRPELKEFTL
 HDHVEAQAFRLQVNGSQPSIFGVHVRGDIYVMPKVKVGVHVRDYLEKALD
 RFRARYSPVFTVSNMKNRENTSLGVDVFNAGVIEGSPAKDFALLTQCNTHTM
 TIGFTGWAAYLAGDITLYLANVLPDPSPLKIFKPAALPEWM"

BASE COUNT
 214 a 285 c 257 g 207 t
 ORIGIN
 Query Match 71.6%; Score 764.6; DB 10; Length 963;
 Best Local Similarity 89.9%; Pred. No. 2.2e-190;

Matches 850; Conservative 0; Mismatches 74; Indels 2; Gaps 2;
 QY 1 CTCACAGCGAATAGTGAAGCTCCAAACCCCTGTTCAGAGAAAGAAATTAACGATGACGACT 60
 Db 40 CTCACAGCAACAAATAGTGAAGCTCCAAACCCCTGTTCAGAGAAAGAAATTAAC---AGGCGGT 96
 QY 61 CAAATGTCTCTGGGAAACACAGAAAGCCAGAGATGCGACGGACGAGCAGCAGCATGGG 120
 Db 97 CAAATGTCTCTCACCAACGCGGCAAGAACAGACATGCGACGCA-----G 138
 QY 121 AATGAGAGAGCTGCGGGGATGTTTACGATCAATTCATTTGGCGGCTGGGAACAGATG 180
 Db 139 AGTGCAAGCTGACGGGATATTCAGATCAATTCATTCGCGGCTGGGAACAGATG 198
 QY 181 GCGAATACGACACACTCTTTGCACTGCGCAGGATGACGACGCGCTTGGCTTCATCCGC 240
 Db 199 GCGAATATGCTACATTTGTTGCACTGCGCAGGATGACGCGCTTGGCTTCATCCGC 258
 QY 241 GCATCCATGCACAAAGCTCTAGCGGCCATCTTCAGGATCAGCTCCCGTGTTCACACAGC 300
 Db 259 GAATCCATGCACAGCTCTAGCGGCCATCTTCAGGATCAGCTCCCGTGTTCACACAGC 318
 QY 301 GACAGCGCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGT 360
 Db 319 GACAGCGCAAGAGATCCCGTGGCAGAAATACCATCTCAACGACTGGATGGAGAGCGT 378
 QY 361 TACCGCCACATTCGCGGACACTTTTGGCTTTCAGGATACCCGCTGCTTCGACCTTC 420
 Db 379 TACCGCCACATTCGCGGAGCAGTATGCTGCTTTCAGGATACCCGCTGCTTCGACCTTC 438
 QY 421 TACCACCACTTCGCGCCAGAGATCCCTGAAGAGTTCACCTTCATGACACGCTGGGAG 480
 Db 439 TACCACCACTTCGCGCCAGAGATCCCTGAAGAGTTCACCTTCAGGATACCCGCTGCTTC 498
 QY 481 GAGCGCCAGGCTTCCTCGCTGGTGTGCGGGTGAATGGAGCCAGGAGTCTTTGTG 540
 Db 499 GAGCGCCAGGCTTCCTCGCTGGTGTGCGGGTGAATGGAGCCAGGAGTCTTTGTG 558
 QY 541 GGTGCTCATGTGCGCGAGGAGCTATGCTGATGTCATGCTGCTGCTGCTGCTGCTGCTG 600
 Db 559 GGTGCTCATGTGCGCGAGGAGCTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 618
 QY 601 GTGCTGACCGGGTTCCTGGAAGAGCCCTGGATATGTTCCGGGACGCTATCACT 660
 Db 619 GTGCTGACCGGGTTCCTGGAAGAGCCCTGGAGAGTTCACCTTCAGGATACCCGCTGCTTC 678
 QY 661 CAGTCTTCGTGTTACAAGCAACGGTATGGCTTGGTCCGGGAGAACATTAATGCTTCC 720
 Db 679 CAGTCTTCGTGTTACAAGCAACGGTATGGCTTGGTCCGGGAGAACATTAATGCTTCC 738
 QY 721 CGAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTCCGACGCAAGGACTTCGCG 780
 Db 739 CTAGGAGACGTGGTGTTCGCGGCAATGGTATTGAGGGGTCCGACGCAAGGACTTCGCG 798
 QY 781 CTGCTCACCAGTGCACACACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 840
 Db 799 CTCCTCACCAGTGCACACACCATCATGACTATTGGGACTTTGGGATTTGGGCTGCC 858
 QY 841 TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTTCCTCGTTC 900
 Db 859 TACCTGGCAGTGGTGATACCATCTACTTAGCCAACTACACCTTCGCGATTTCCTCGTTC 918
 QY 901 CTCAAAGTCTTTAAGCCAGAGGACGCTTCCTACCCGGAATGGGTG 945
 Db 919 CTCAAAATCTTTAAGCCAGAGGACGCTTCCTCCCGAGTGGATG 963

RESULT 11
 AB039116
 LOCUS
 DEFINITION
 Mus musculus Fut2 gene for GDP-L-fucose:beta-D-galactoside
 2-alpha-L-fucosyltransferase, partial cds, strain: B1G2/Msf.
 AB039116
 ACCESSION

```

AB039116.1 GI:15822977
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
Mus musculus (sub_species:musculus, strain:BLG2/Msf) DNA.
Mus musculus
ORGANISM
REFERENCE
AUTHORS
Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
Unpublished
JOURNAL
REFERENCE
AUTHORS
Liu,Y. and Saitou,N.
Direct Submission
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="BLG2/Msf"
/sub_species="musculus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1..963
/gene="FUT2"
<1..>963
/gene="FUT2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase"
/protein_id="BAB68640.1"
/db_xref="GI:15822978"
/translation="ELIFVFVISTIIHLOORIVKLTSEKELQAVOMSSPNAFTDM
QOSALQGIITNSIGLNGNGEYATLEFALAMNGRLAFIPESMENALAFIRSLP
VLHSTARIIPQVHLNDWMEERHPIQVRYRTGYPCSWTFYHLRPEILKEFTL
HDHVRPEAQFLRLGVNSQPSFVGVVRRGDYVHVMPKVGKVVADRGYLEKALD
REFRYSFVVFVTSNGMAWCNENTSLGDVVFAGNGEGSPAKDFALLTCNHTIM
TIGTGIWAIYLAGDITLYLANTLPDSFFLKIFPAALFPEWM"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative. 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCACACAGGATATGAGCTCCAAACCCCTGTCAGAGGAAGAAATACGATGACGACT 60
Db 40 CTCACACAGGATATGAGCTCCAAACCCCTGTCAGAGGAAGAAATAC ---AGCGGTT 96
QY 61 CAATGCTCTCGGGAACACAGAAAGCCAGAGATCGGACGGACAGCGAGCATGGG 120
Db 97 CAATGCTCTCACCACCGGCGGAGAGACAGATCGACA-----G 138
QY 121 AATGGAGAGCTGGGGGCGATGTTTCAGATCAATTCATTGGCGGCTGGGGACACAGATG 180
Db 139 AGTGCCAAGCTGAGGGCATATTCAGATCAATTCATCGCGCTGGGGAACACAGATG 198
QY 181 GCGATATGCGACATCTTTTCAGTGGCGAGATGAACGGAGCGGTTGGGTTTCATCCCC 240
Db 199 GCGAATATGCTACATGTTTGCACGTGGCGAGATGAACGGTGGCTTCCTTCCTCC 258
QY 241 GCATCCATGCACAAACGCTCTAGCGCCCATCTTCAGATCAGCTCCCGGTGTTACACAGC 300
Db 259 GAATCCATGCACAAACGCTCTAGCGCCCATCTTCAGATCAGCTCCCGGTGTTACACAGC 318

```

```

QY 301 GACACGCCCAAAAGATCCATGGCAGAAATACCATCTCAACGACTGGAGGAGCGT 360
Db 319 GACACGCCCAAGAGATCCGTCGAGAGATACCACTCAACGACTGGAGGAGCGT 378
QY 361 TACCGCCCATTCGCGGACACATTTGTGCGCTTACCGGATACCGGCTCTCTCGACTTC 420
Db 379 TACCGCCCATTCGCGGCGAGTATGTGCGTTTACCGGATACCGGCTCTCTCGACTTC 438
QY 421 TACCACCACTCGGCCCGCCAGAGATCCTGAAGGAGTTCACCTGCATCACCAGCTGGCGGAG 480
Db 439 TACCACCACTCGGCCCGCCAGAGATCCTGAAGGAGTTCACCTGCATCACCAGCTGGCGGAG 498
QY 481 GAGGCCAGCGCTTCTCGGCTGGTCTCGGCTGAATGGGAGCCAGCCGAGTACTTTGTG 540
Db 499 GAGGCCAGCGCTTCTCGGCTGGTCTCGGCTGAATGGGAGCCAGCCGAGTACTTTGTG 558
QY 541 GGTGTCCATGTGCGCGCGAGGAGACTATGTGCATGTATGCTCATGCTGGAAGGCGGTG 600
Db 559 GGTGTCCATGTGCGCGCGAGGAGACTATGTGCATGTATGCTCATGCTGGAAGGCGGTG 618
QY 601 GTGGCTGACCGGGTTACCTGGAAAGGCCCTGGATATGTCGGGCGACGCTATTCATCT 660
Db 619 GTGGCTGACCGGGTTACCTGGAAAGGCCCTGGAGAGTTCGGGCGACGCTATTCATCT 678
QY 661 CCAGTCTTCGTGTTTACAAGCAACGGTATGGCTGGTCCGGGAGAACATTAIGCTCC 720
Db 679 CCAGTCTTCGTGTTTACAAGCAACGGTATGGCTGGTCCGGGAGAACATCAACACTCC 738
QY 721 CGAGGAGAGCTGTGTTCGCGGCAATGGTATGAGGGTTCGCCAGCCAGGACTTCGGG 780
Db 739 CTAGGAGAGCTGTGTTCGCGGCAATGGTATGAGGGTTCGCCAGCCAGGACTTCGGG 798
QY 781 CTGCTCACCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGC 840
Db 799 CTCTCACCAGTGCACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGC 858
QY 841 TACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTACACCTTCGGGATTCGGTTC 900
Db 859 TACCTGGCAGGTGGTGATACCATCTACTTAGCCAACTACACCTTCGGGATTCGGTTC 918
QY 901 CTCAAAGCTTTAAGCCAGAGGAGCGCTTCCTACCCGAGTGGGTG 945
Db 919 CTCAAAGCTTTAAGCCAGAGGAGCGCTTCCTCCCGAGTGGATG 963
RESULT 12
AB039117
LOCUS
DEFINITION
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-1-fucosyltransferase, partial cds, strain:CAST/Ei.
AB039117
ACCESSION
AB039117.1 GI:15822979
VERSION
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-1-fucosyltransferase.
Mus musculus (sub_species:castaneus, strain:CAST/Ei) DNA.
KEYWORDS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ORGANISM
REFERENCE
1
AUTHORS
Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
Unpublished
JOURNAL
REFERENCE
AUTHORS
Liu,Y. and Saitou,N.
Direct Submission
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
source

```

/organism="Mus musculus"
/strain="CAST/Ei"
/sub_species="castaneus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1..963
/gene="FUT2"
<1..>963
/gene="FUT2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB68641.1"
/db_xref="GI:15822980"

gene
CDS

QOSAKLOGIETINSIGRLGNQMGYATLIFALARNGLAFIPESMHNALAFIPESLIP
VLHSDARRIPWONHLDNMEERYRHPGQVRFYPCSWTFYHHRPEILKEFTL
HDHVREAAQAFRLGLRVNGSOPSTFVGVHVRGQYVHVMPKVMGVVADRGYLEKALD
RPRARYSPFVFTVSNMAHRENIINTSLGDVFAGNIEGSPAKDFALLIQCNHTIM
TIGTFGLWAAYLAGDITIIYLANITLIPSPFPIKIPAAFLPEWH"
BASE COUNT 214 a 285 c 257 g 207 t

Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCAGCAGCGAATAGTGAAGCTCCAAACCCCTGTGACAGAGAGGAATTACCGATGACGACT 60
DB 40 CTCAGCAGCAAGATAGTGAAGCTCCAAACCCCTGTGACAGAGAGGAATTAC---AGCGGTT 96
QY 61 CAATGCTCTCGGNAACACAGAACCCACAGAGATGCGACGGGACGACGAGCATGGG 120
DB 97 CAATGCTCTCACCNAACGCGGACAGACATGCAGCA-----G 138
QY 121 AATGAGAGCTCGGGGATGTTCAGATCAATTCATTCGCGGCTGGGGAACAGAGAT 180
DB 139 AGTGCAAGCTGACGGGATATTCAGATCAATTCATTCGCGGCTGGGGAACAGAGAT 198
QY 181 GCGAATACGCCACACTCTTTCAGTCCGCGGATGAACGAGCGCTTGCCTTCATCC 240
DB 199 GCGAATATGCTACATCTTTCAGTCCGCGGATGAACGAGCGCTTGCCTTCATCC 258
QY 241 GCATCCATGCACAGCTCTAGCGGCCATCTTCAGATCAGCTCCCGGTGTTACACAGC 300
DB 259 GAATCCATGCACAGCTCTAGCGGCCATCTTCAGATCAGCTCCCGGTGTTACACAGC 318
QY 301 GACAGCGCAAAAGATCCCATGCGAGAAATACCATCTCAAGCATGATGGAGGAGCGT 360
DB 319 GACACAGCGCAGAGATCCCGTGGCAGAAATACCACTCAAGCATGATGGAGGAGCGT 378
QY 361 TACCGCCACATTCGCGGACACTTGTGCGCTTCACGGATACCCGTGCTTCGACCTTC 420
DB 379 TACCGCCACATTCGCGGACACTTGTGCGCTTCACGGATACCCGTGCTTCGACCTTC 438
QY 421 TACACACAGCTGCGCCACAGATCTGAAGAGTTCACCTGCATGACACAGTGGGGAG 480
DB 439 TACACACAGCTGCGCCACAGATCTGAAGAGTTCACCTGCATGACACAGTGGGGAG 498
QY 481 GAGGCCAGGCTCTCTGCGTGGTCTGCGGGTGAATGGGAGCAGCGAGTACTTTTGTG 540
DB 499 GAGGCCAGGCTCTCTGCGTGGTCTGCGGGTGAATGGGAGCAGCGAGTACTTTTGTG 558
QY 541 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTTCATGCTCATGCTGTGGAGGCGCTG 600
DB 559 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTTCATGCTCATGCTGTGGAGGCGCTG 618
QY 601 GTGGCTGACCGGGGTACTGTGAAAAGGCCCTTGGATATGTTCGCGGACGCTATTCACT 660
|||||

DB 619 GTGCTGACCGGGTTACCTAGAAAAGCCCTGGACAGGTTCCGGGCACGCTATTTCATCT 678
QY 661 CCAGCTTCGTTGTTACAGCAACGTTATGGCTGTCGGGAGAACATTAATGCTTCC 720
DB 679 CCAGCTTCGTTGTTACAGCAACGTTATGGCTGTCGGGAGAACATTAATGCTTCC 738
QY 721 CGAGGAGACGTTGTTGTCGGGCAATGTTAGGGGTCGCCAGCCAAAGCACTTCGGG 780
DB 739 CTAGGAGACGTTGTTGTCGGGCAATGTTAGGGGTCGCCAGCCAAAGCACTTCGGG 798
QY 781 CTGCTACCCAGTGCACACACACCATCATGATATTTGGACCTTTGGGATTTGGCTGCC 840
DB 799 CTCTCACCCAGTGCACACACACCATCATGATATTTGGACCTTTGGGATTTGGCTGCC 858
QY 841 TACCTGGCAGTGTGATACCATCTACTTAGCAACTACACCTTCCCGGATTTCTCGTTC 900
DB 859 TACCTGGCAGTGTGATACCATCTACTTAGCAACTACACCTTCCCGGATTTCTCGTTC 918
QY 901 CTCAAAGTCTTAAGCCAGCAGCAGCTTCCCTACCCGAATGGTG 945
DB 919 CTCAAAGTCTTAAGCCAGCAGCAGCAGCTTCCCTCCCGAGTGGATG 963

RESULT 13

AB039118
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

AB039118 963 bp DNA linear ROD 02-OCT-2001
Mus musculus FUT2 gene for GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase, partial cds, strain:MSM/Msf.
AB039118.1 GI:15822981
FUT2; GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.
Mus musculus (sub_species:molesinus, strain:MSM/Msf) DNA.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
1
Liu,Y., Kitano,T., Koide,T., Shiroishi,T., Moriaki,K. and
Saitou,N.
Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes
of Five Mus musculus subspecies
unpublished
2 (bases 1 to 963)
Liu,Y. and Saitou,N.
Direct submission
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayerlab.nig.ac.jp/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="MSM/Msf"
/sub_species="molesinus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1..963
/gene="FUT2"
<1..>963
/gene="FUT2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB68641.1"
/db_xref="GI:15822982"
QOSAKLOGIETINSIGRLGNQMGYATLIFALARNGLAFIPESMHNALAFIPESLIP
VLHSDARRIPWONHLDNMEERYRHPGQVRFYPCSWTFYHHRPEILKEFTL
HDHVREAAQAFRLGLRVNGSOPSTFVGVHVRGQYVHVMPKVMGVVADRGYLEKALD

[illegible]

```
199 GCGAATATGCTACATTTGCTAGCTGGCCAGGATGACGGTGGCTTCATCCCT 258
241 GCATCCATGACAAAGCTCTAGCGCCCATCTTACAGATCAGCTCCCGTGTACACAGC 300
259 GAATCCATGACAAAGCTCTAGCGCCCATCTTACAGATCAGCTCCCGTGTACACAGC 318
301 GACAGGCCAAAGATCCCATGCGAGATTAACATCTCACAGCTGGATGAGAGCGT 360
319 GACAGCCAGAGATCCCGTGGCAGATTAACCATCTCACAGCTGGATGAGAGCGT 378
361 TACCGCCACATTCGGGACACTTTGTGGCTTACAGGATACCGGTGCTCTGGACCTTC 420
379 TACCGCCACATTCGGGACACTTTGTGGCTTACAGGATACCGGTGCTCTGGACCTTC 438
421 TACACACCTTGGCCCGCAGAGATCTGAGAGAGTTCACCTGCTGATGACCGTGGGAG 480
439 TACCACACCTTGGCCCGCAGAGATCTGAGAGAGTTCACCTGCTGATGACCGTGGG 498
481 GAGGCCAGGCTTCCCTGGTGGTCTGGGCTGATGGAGCCAGCCAGTACTTTGTG 540
499 GAGGCCAGGCTTCCCTGGTGGTCTGGGCTGATGGAGCCAGCCAGTACTTTGTG 558
541 GGTGTCCATGTCCCGCAGGAGTATGTGATGTGATGCTGATGCTGATGCTGATGCTG 600
559 GGTGTCCATGTCCCGCAGGAGTATGTGATGTGATGCTGATGCTGATGCTGATGCTG 618
601 GTGGTACCGGGTACCTGGAAAGCCCTGATATGTTCCGGGACAGCTATTCATCT 660
619 GTGGTACCGGGTACCTGGAAAGCCCTGATATGTTCCGGGACAGCTATTCATCT 678
661 CCAGTCTTGGTGTACAGCAAGGATGCTGCTGGTGGTGGGAGAACATTAATGCTTCC 720
679 CCAGTCTTGGTGTACAGCAAGGATGCTGCTGGTGGTGGGAGAACATTAATGCTTCC 738
721 CGAGGAGCGTGGTGTCCCGGCAATGATGATGAGGGTGGCCGCAAGGACTTCGCG 780
739 CTAGGAGCGTGGTGTCCCGGCAATGATGATGAGGGTGGCCGCAAGGACTTCGCG 798
781 CTGCTACCGAGTGAACCAACACATCATGATATGAGGAGTGGAGTGGGATTCGCTCC 840
799 CTCCTACCGAGTGAACCAACACATCATGATGATGAGGAGTGGAGTGGGATTCGCTCC 858
841 TACCTGGAGTGGTGTACCATCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 900
859 TACCTGGAGTGGTGTACCATCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 918
901 CTCAAAGTCTTAAAGCCAGAGGAGCTTCTCTACCGGAATGGTG 945
919 CTCAAAGTCTTAAAGCCAGAGGAGCTTCTCTCCCGGAGTGATG 963

RESULT 15
AB039120
LOCUS
DEFINITION
AB039120
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
```

```
JOURNAL
Submitted (29-FEB-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genetics.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/, Tel:81-559-81-6730,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..963
/organism="Mus musculus"
/strain="SWN/Msf"
/sub_species="molossinus"
/db_xref="taxon:10090"
/chromosome="7"
/note="sequence used for primer design based on
Acc#AF064792"
1..963
/gene="PUT2"
<1..>963
/gene="PUT2"
/EC_number="2.4.1.69"
/note="Sec2; Se"
/codon_start=1
/product="GDP-L-fucose:beta-D-galactoside
2-alpha-L-fucosyltransferase"
/protein_id="BAB68644.1"
/db_xref="GI:15822986"
/translation="FLIPVFVITSTIHLQQRIVKLOTLSKELQKVMSSPNAARDIM
QOSAKLOGIFTINSTIGLGNOMGEYATIFALARMNGLAFIPESHNHNLAFIPRISLP
VLHSDARRIPWQNYHLNDWEERYRHIPQGYRFTGYPCSWTFYHHLRPIILKEFTL
HDVREAAQAFRLGLRVNQSPTFVGVHVRGDDVHVMPKVGKADRGYDLKALD
RFRYRSPFVFTVSTNGMARENITSLDGVFAGNIPKSPKVDKADFTLTCNNHTIM
TIGTFGIWAAYLAGDITILANTYLPDPSFLKIFKPAALFPEW"
BASE COUNT 214 a 285 c 257 g 207 t
ORIGIN
Query Match 71.6%; Score 764.6; DB 10; Length 963;
Best Local Similarity 89.9%; Pred. No. 2.2e-190;
Matches 850; Conservative 0; Mismatches 74; Indels 21; Gaps 2;
QY 1 CTCGAGCAGCAATAGTGAAGCTCCAAACCCCTGCAGAGAGGAATTACCGATGACGACT 60
DB 40 CTCGAGCAGCAATAGTGAAGCTCCAAACCCCTGCAGAGAGGAATAG---AGCGGTT 96
QY 61 CAAATGTCTCTGGGAAACACAGAAAGCCGACGAGATCCGACGGGACAGCAGCATGGG 120
DB 97 CAAATGTCTCTACCAACACGCGGCAAGACAGACATGACGCA-----G 138
QY 121 AATGAGAGCTCGGGGATGTTACAGATCAATTCATTCGCGCGCTGGGACAGATG 180
DB 139 AGTGCCAAAGCTCGAGGCAATATTCAGATCAATTCATTCGCGCGCTGGGAAACAGATG 198
QY 181 GCGCAATAGCCACACTCTTTCACCTGGCCAGGATGAACGGACGGCTTCGCTTCATCCCC 240
DB 199 GCGCAATAGCTACATTTGTTGCACTGGCCAGATGAACGGTTCGCTTCATCCCT 258
QY 241 GCATCCATGACAAACGCTCTAGCGCCATCTTTCAGATCAGCTCCCGTGTTCACACAGC 300
DB 259 GAATCCATGACAAACGCTCTAGCGCCATCTTTCAGATCAGCTCCCGTGTTCACACAGC 318
QY 301 GACAGGCCAAAGATCCCATGGCAGATTAACATCTCACAGCTGGATGAGAGCGT 360
DB 319 GACAGCCAGAGATCCCGTGGCAGATTAACCATCTCACAGCTGGATGAGAGCGT 378
QY 361 TACCGCCACATTCGGGACACTTTGTGGCTTACAGGATACCGGTGCTCTGGACCTTC 420
DB 379 TACCGCCACATTCGGGACACTTTGTGGCTTACAGGATACCGGTGCTCTGGACCTTC 438
QY 421 TACCACACCTTGGCCCGCAGAGATCTTGAAGAGTTCACCTGATGACCATCTGGGAG 480
DB 439 TACCACACCTTGGCCCGCAGAGATCTTGAAGAGTTCACCTGATGACCATCTGGTGA 498
QY 481 GAGGCCAGGCTTCCCTGGTGGTCTGGGCTGATGGAGCCAGCCAGTACTTTGTG 540
DB 499 GAGGCCAGGCTTCCCTGGTGGTCTGGGCTGATGGAGCCAGCCAGTACTTTGTG 558
```

```
FEATURES
source
```

```
QY 541 GGTGTCATGTGCCCGAGGGGACTATGTGTCATGTGCTATGCTAATGTGIGGAGGGGCTG 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 559 GGTGTCATGTGCCCGAGGGGACTATGTGTCATGTGCTAATGCTGAGGGGTGGAAGGGCTG 618
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 GTGGCTGACCGGGGTTACCTGGAAAGGCCCTGGATATGTTCCGGGCACGCTATTTCATCT 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 619 GTGGCTGACCGGGGTTACCTAGAAAGGCCCTGGACAGGTTCCGGGCACGCTATTTCATCT 678
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 CCACTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTCCCGGGAGACANTTAATGCTTCC 720
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 679 CCACTCTTCGTGGTTACAAGCAACGGTATGGCCTGGTCCCGGGAGACATCAACACCTCC 738
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 721 CGAGGAGACGTGGTGTTCGCGGGCAATGGTATTGAGGGGTGCGCAGCCCAAGACTTGGCG 780
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 739 CTAGGAGACGTGGTGTTCGCGGGCAATGGTATTGAGGGGTGCGCAGCCCAAGACTTGGCG 798
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 781 CTGCTCACCCAGTGCACACCAACCATCATGACTATTGGGACCTTTGGGATTGGGCTGCC 840
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 799 CTCCTCACCCAGTGCACACCAACCATCATGACTATTGGGACCTTTGGGATTGGGCTGCC 858
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 841 TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCCCTCCGGATTCTCCGTTTC 900
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 859 TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCCCTCCGGATTCTCCGTTTC 918
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 901 CTCRAAGTCTTTAAGCCAGAGGCGCCCTTCCTACCCGAATGGGTG 945
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 919 CTCRAATCTTTAAGCCAGAGGCGCCCTTCCTACCCGAATGGGTG 963
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: May 26, 2003, 14:49:17
Job time : 2883.28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 09:17:27 ; Search time 276.996 Seconds
(without alignments)
8882.922 Million cell updates/sec

Title: US-10-040-863-9

Perfect score: 1068

Sequence: 1 ctccagcagcgaatagttaa.....gaagagccttctgatggagg 1068

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068	100.0	1068	21 AAC67966	Rat hepatoma H35 c
2	1068	100.0	1068	24 AAD27208	Rat hepatoma H35 c
3	1068	100.0	1149	21 AAC67965	Rat hepatoma H35 c
4	1068	100.0	1149	24 AAD27207	Rat hepatoma H35 c
5	612.8	57.4	1043	19 AAV29003	Porcine secretor t
6	612.8	57.4	1043	19 AAV21639	Pig secretor encod
7	589.4	55.3	2115	19 AAV38323	Human Sec2 coding
8	589.4	55.2	3088	24 ABL64735	Lung cancer relate
9	589.4	55.2	3088	24 ABL66311	Lung cancer relate

10	589.4	55.2	3088	24	ABL70026	Pancreas cancer re
11	423.8	39.7	1098	19	AAV21640	Pig H transferase
12	423.8	39.7	1269	20	AAV15872	Fragment of pHR3x
13	423.8	39.7	1269	20	AAV03811	Swine alpha-1,2-fu
14	423.8	39.7	1269	21	AAZ94417	Pig alpha-1,2 fuco
15	423.8	39.7	1269	24	AAI72831	FU2 cDNA. Sus sc
16	405	37.9	1155	16	AAQ98461	GDP-L-fucose-beta-
17	405	37.9	1155	16	AAQ01082	2-Alpha-fucosyltra
18	405	37.9	1155	16	AAQ01083	Human H-transferas
19	405	37.9	1174	17	AAI12238	Human H-transferas
20	405	37.9	1174	21	AAV53820	Mouse H2kb gene pr
21	405	37.9	2268	18	AAV63576	Human alpha 1,2 fu
22	405	37.9	3733	18	AAV67678	Chicken beta-actin
23	405	37.9	3791	18	AAV63575	GDP-Fuc:beta-D-gal
24	405	37.9	8174	12	AAQ13332	Human alpha(1,2)-f
25	405	37.9	8174	18	AAV61677	DNA encoding a g-y
26	403.4	37.8	8174	15	AAQ56908	X. laevis alpha1,
27	304.8	28.5	1071	20	AAV7268	Human secreted pro
28	125.4	11.7	473	21	AAV03536	Rabbit alpha1-2fuc
29	72.8	6.8	100	21	AAV67984	Rabbit alpha1-2fuc
30	71.2	6.7	100	21	AAV67978	Rabbit alpha1-2fuc
31	71.2	6.7	100	21	AAV67981	Rabbit alpha1-2fuc
32	69.6	6.5	100	21	AAV67975	Human alpha1-2fuco
33	63.2	5.9	100	21	AAV67969	Human alpha1-2fuco
34	63.2	5.9	100	21	AAV67972	Human alpha1-2fuco
35	44	4.1	495	22	AAH06199	Human cDNA clone (
36	42.8	4.0	4069	22	AAH17772	Human cDNA sequenc
37	42.8	4.0	5236	22	AAH16566	Human polynucleoti
38	42.8	4.0	5530	22	AAK52640	Human polynucleoti
39	40.4	3.8	2297	22	AAV05087	Human secreted pro
40	40.4	3.8	2612	22	AAV05059	Human secreted pro
41	40.4	3.8	2953	23	ABK43445	DNA encoding novel
42	40.4	3.8	3044	22	ABK31284	Human cDNA encodin
43	40.4	3.8	3044	24	ABO66608	Human polynucleoti
44	38.4	3.6	50	21	AAV67971	Human alpha1-2fuco
45	38.2	3.6	51	22	AAV79569	Human DNA containi

ALIGNMENTS

RESULT 1

AAAC67966

ID AAC67966 standard; cDNA; 1068 BP.

XX AC AAC67966;

XX AC AAC67966;

XX DT 19-FEB-2001 (first entry)

XX DE Rat hepatoma H35 cell alpha1-2fucosyltransferase catalytic domain cDNA.

XX DE Rat; alpha1-2fucosyltransferase; cytosolic; neuroprotective;

XX KW nototropic; gene therapy; Fucalalpha1-2Galbeta1-3GalNAc; immunotherapy;

XX KW immunosuppression; cancer; neurological disease;

XX KW small cell lung carcinoma; ss.

XX OS Rattus norvegicus.

XX OS Rattus norvegicus.

XX PN WO2000064464-A1.

XX XX 02-NOV-2000.

XX XX 23-APR-1999; 99WO-US07384.

XX XX 23-APR-1999; 99WO-US07384.

XX XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.

XX PI Holmes EH, Sherwood AL;

XX XX WPI: 2000-687262/67.

XX DR P-PSDB; AAB36105.

XX XX

PT New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for
PT preparation of fucosyl GM1 which is useful as a nutritional composition
XX or immunotherapeutic for cancer and neurological diseases -
XX Claim 11; Fig 3A; 91pp; English.

XX The present sequence is given in a specification relating to a rat
CC ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein
CC or its cellular fraction is useful for synthesis of a molecule comprising
CC Fucalpal-2galbetar-3galNAc, a glycolipid, glycoprotein, glycolipoprotein
CC or a free oligosaccharide comprising Fucalpal-2galbetar-3galNAc.
CC The method involves contacting alpha1-2fucosyltransferase with GDP-fucose
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or
CC oligosaccharide having a terminal Galbetar-3galNAc group. It is also
CC useful for synthesis of fucosyl-GM1 by contacting the protein with
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
CC glycolipoproteins, glycolipids and oligosaccharides are useful as
CC nutritional compositions and fucosyl-GM1 is useful for inducing an
CC immunotherapeutic or immunosuppressive action against cancer,
CC neurological disease or small cell lung carcinoma.

SQ Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 100.0%; Score 1068; DB 21; Length 1068;

Best Local Similarity 100.0%; Pred. No. 8.5e-289;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGACGAGTATAGTACGCTCAACCCCTGTGACGAGAGGAATACCGATGACGACT 60
DB 1 CTCACGACGAGTATAGTACGCTCAACCCCTGTGACGAGAGGAATACCGATGACGACT 60
QY 61 CAATGTCTCGGGGAAACACAGAAAGCCAGAGATCGACGGGACGAGCAGCAGATGG 120
DB 61 CAATGTCTCGGGGAAACACAGAAAGCCAGAGATCGACGGGACGAGCAGCAGATGG 120
QY 121 AATGAGAGCTCGGGGATGTTACGATCAATTCATTGGCCGCTGGGGAACACGATG 180
DB 121 AATGAGAGCTCGGGGATGTTACGATCAATTCATTGGCCGCTGGGGAACACGATG 180
QY 181 GCGAATAGCCACACTCTTTCAGTCCAGGATGAACGGAGCGGTTCCTATCC 240
DB 181 GCGAATAGCCACACTCTTTCAGTCCAGGATGAACGGAGCGGTTCCTATCC 240
QY 241 GCATCCATGCACACGCTTACGCGCCATCTCAGATCAGCTCCCGTGTATACACAG 300
DB 241 GCATCCATGCACACGCTTACGCGCCATCTCAGATCAGCTCCCGTGTATACACAG 300
QY 301 GACACGGCCAAAAGATCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT 360
DB 301 GACACGGCCAAAAGATCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT 360
QY 361 TACCGCCACATTCGGGACACTTTGCGCTTCACGGGATACCGTCCCTGGACCTTC 420
DB 361 TACCGCCACATTCGGGACACTTTGCGCTTCACGGGATACCGTCCCTGGACCTTC 420
QY 421 TACCACCACTCGGCCAGAGATCTCTGAAGAGATTCACCTGCATCACCACGCTGGCGGAG 480
DB 421 TACCACCACTCGGCCAGAGATCTCTGAAGAGATTCACCTGCATCACCACGCTGGCGGAG 480
QY 481 GAGGCCAGCCCTCTCTGCTGGTCTGCGGGTGAATGGAGGACGACCGAGTCTTTTGTG 540
DB 481 GAGGCCAGCCCTCTCTGCTGGTCTGCGGGTGAATGGAGGACGACCGAGTCTTTTGTG 540
QY 541 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTCATGCCTATGTGTGGAGGCGGTG 600
DB 541 GGTGTCCATGTGCGCGAGGGGACTATGTGCATGTCATGCCTATGTGTGGAGGCGGTG 600
QY 601 GTGGCTACCGGGTTCACCTGGAAGAGCCCTGGATATGTTCGGGACGCTATTCATCT 660
DB 601 GTGGCTACCGGGTTCACCTGGAAGAGCCCTGGATATGTTCGGGACGCTATTCATCT 660
QY 661 CCAGTCTCTGTGTTTACAAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720
DB 661 CCAGTCTCTGTGTTTACAAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720

DB 661 CCAGTCTCTGTGTTTACAAGCAACGATGGCTGTGCGGGAGAACATTAATGCTTCC 720
QY 721 CGAGAGACSTGGTGTTCGCGGGCAATGTTAGGGGTCCCGAGCAAGACTTCGGG 780
DB 721 CGAGAGACSTGGTGTTCGCGGGCAATGTTAGGGGTCCCGAGCAAGACTTCGGG 780
QY 781 CTGCTFACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGTC 840
DB 781 CTGCTFACCCAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGTC 840
QY 841 TACCTGGCAGGGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCGGTTC 900
DB 841 TACCTGGCAGGGGTGATACCATCTACTTAGCCAACTACACCTTCCGGATTCGGTTC 900
QY 901 CTCGAAGTCTTTAAGCCAGAGCAGCTTCTACCCGAATGGTGGGATCCCTGCCGAT 960
DB 901 CTCGAAGTCTTTAAGCCAGAGCAGCTTCTACCCGAATGGTGGGATCCCTGCCGAT 960
QY 961 CTGTCCCACTCTTAAAGGCATTAACACAGCAGCTGTCTCTGGTCCGCTCCACCTCAAG 1020
DB 961 CTGTCCCACTCTTAAAGGCATTAACACAGCAGCTGTCTCTGGTCCGCTCCACCTCAAG 1020
QY 1021 GCATAAGAGTCACTTGTACCTCCGAGCAAGAGCTTCTGTGATGGAA 1068
DB 1021 GCATAAGAGTCACTTGTACCTCCGAGCAAGAGCTTCTGTGATGGAA 1068
RESULT 2
AAD27208
ID AAD27208 standard; DNA; 1068 BP.
XX AAD27208;
XX AC
XX DT
XX 09-APR-2002 (first entry)
XX Rat hepatoma H35 cell alpha1-2FucT catalytic domain encoding DNA.
XX Rat; alpha1-2fucosyltransferase; alpha1-2FucT; antisense therapy;
KW galactose beta1-3N-acetylglactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
KW cell transformation; catalytic domain; ds.
XX Rattus norvegicus.
XX OS
XX FH
XX FT
FT CDS
FT Location/Qualifiers
FT 1..1062
FT /tag= a
FT /product= "Rat hepatoma H35 cell alpha1-2FucT catalytic
FT domain"
FT /note= "CDS does not include start codon"
FT /partial
FT 531..1068
FT misc_feature
FT /tag= b
FT /note= "Region which overlaps rat FTB"
XX US6329170-B1.
XX PN
XX PD
XX 11-DEC-2001.
XX XX
XX 23-APR-1999; 99US-0298886.
XX PF
XX 23-APR-1999; 99US-0298886.
XX PR
XX (NWHO-) NORTHWEST HOSPITAL.
XX PA
XX Holmes EH, Sherwood AL;
XX WPI; 2002-121132/16.
XX P-PSDB; AAE16623.
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant

techniques and for detecting oncogenic transformation of test tissues -

Claim 2; Fig 3; 41pp; English.

The invention relates to rat GM1-specific alphas-2-fucosyltransferase (alpha-2-fucT) enzyme and its corresponding nucleic acid. This nucleic acid is specific for a carbohydrate moiety found in ganglioside GM1, a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc) saccharide. Alpha-2-fucT DNA is useful for producing rat alpha-2-fucT protein by recombinant techniques. Alpha-2-fucT DNA is useful for the preparative synthesis of fucosyl containing glycolipids, glycoproteins, CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1. CC Alpha-2-fucT DNA is useful for detecting oncogenic transformation which CC involves assaying for changes in expression of alpha-2 fucT. Since CC alpha-2-fucT is activated in cell transformation, antisense sequences CC derived from alpha-2-fucT DNA are useful for inhibiting, suppressing CC or treating cancer. Alpha-2-fucT DNA is useful in gene therapy and CC antisense therapy. The present sequence is rat hepatoma H35 cell CC alpha-2-fucT catalytic domain encoding DNA.

SQ Sequence 1068 BP; 237 A; 308 C; 300 G; 223 T; 0 other;

Query Match 100.0%; Score 1068; DB 24; Length 1068;

Best Local Similarity 100.0%; Pred. No. 8,6e-289;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTCCAGCAGCAATAGTGAAGCTCCAAACCCCTGTGACAGAGAAGTAATACCGATGACGACT	60
DB	1	CTCCAGCAGCAATAGTGAAGCTCCAAACCCCTGTGACAGAGAAGTAATACCGATGACGACT	60
QY	61	CAATGTCCTCGGGAAACACAGAAAGCCAGAGATCGGACGGGACAGCAGCATGCG	120
DB	61	CAATGTCCTCGGGAAACACAGAAAGCCAGAGATCGGACGGGACAGCAGCATGCG	120
QY	121	AATGGAGAGCTCGGGGACATGTTACAGATCAATTCATTGGCCGCTGGGACACAGATG	180
DB	121	AATGGAGAGCTCGGGGACATGTTACAGATCAATTCATTGGCCGCTGGGACACAGATG	180
QY	181	GCGAATACGCCACACTCTTTGACATGGCCAGGATGAACGGGCTTGGCTTCATCCCC	240
DB	181	GCGAATACGCCACACTCTTTGACATGGCCAGGATGAACGGGCTTGGCTTCATCCCC	240
QY	241	GCATCATGCACAAACGCTCTAGCGGCCATCTTCAGGATCAGCCTCCGGTGTACACAG	300
DB	241	GCATCATGCACAAACGCTCTAGCGGCCATCTTCAGGATCAGCCTCCGGTGTACACAG	300
QY	301	GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGT	360
DB	301	GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGCGI	360
QY	361	TACCGCCACATTCGGGACACTTTGCGCTTCACGGGATACCGTCTCCTGGACCTTC	420
DB	361	TACCGCCACATTCGGGACACTTTGCGCTTCACGGGATACCGTCTCCTGGACCTTC	420
QY	421	TACCAACACTGCGGCCACAGATCTCAGAGAGTTACCGTGCATGACACAGTCGGGAG	480
DB	421	TACCAACACTGCGGCCACAGATCTCAGAGAGTTACCGTGCATGACACAGTCGGGAG	480
QY	481	GAGGCCAGGCTCTCTGCGTCTGCGGCTCAANTGGGACGACCGAGTACTTTTGTG	540
DB	481	GAGGCCAGGCTCTCTGCGTCTGCGGCTCAANTGGGACGACCGAGTACTTTTGTG	540
QY	541	GGTGTCCATGTGCGGAGGGGACTATGTCATGTGCTATGCTAAATGTGGAAGGCGGTG	600
DB	541	GGTGTCCATGTGCGGAGGGGACTATGTCATGTGCTATGCTAAATGTGGAAGGCGGTG	600
QY	601	GTGGCTGACCGGGTTACTTGGAAAAGCCCTGGATGTTCGGGACGACCTATTTCATCT	660
DB	601	GTGGCTGACCGGGTTACTTGGAAAAGCCCTGGATGTTCGGGACGACCTATTTCATCT	660
QY	661	CCAGTCTTCTGTTACAGCAGCGTATGGCTTGGTGGCGGAGACATATATGCTTC	720
DB	661	CCAGTCTTCTGTTACAGCAGCGTATGGCTTGGTGGCGGAGACATATATGCTTC	720

QY	721	CGAGGAGAGTGTTGTCGGGCAATGGTATTAGGGGTTCGCCAGCAAGGACTTCGG	780
DB	721	CGAGGAGAGTGTTGTCGGGCAATGGTATTAGGGGTTCGCCAGCAAGGACTTCGG	780
QY	781	CTGCTCACCCAGTGCAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGCC	840
DB	781	CTGCTCACCCAGTGCAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGCTGCC	840
QY	841	TACCTGCGAGGTGTGATACCATCTACTTAGCCAACTACACCTTCGGATTCCTCGCTTC	900
DB	841	TACCTGCGAGGTGTGATACCATCTACTTAGCCAACTACACCTTCGGATTCCTCGCTTC	900
QY	901	CTCAAAAGTCTTTAAGCCAGAGGAGGAGGCTTCTTACCGGAATGGTGGGTCATCCCTGCCGAT	960
DB	901	CTCAAAAGTCTTTAAGCCAGAGGAGGAGGCTTCTTACCGGAATGGTGGGTCATCCCTGCCGAT	960
QY	961	CTGTCGCCACTCTTAAGGCAITTAACACAGAGCTCTCTCGGTCGCCACTTCACCTCAAG	1020
DB	961	CTGTCGCCACTCTCTTAAGGCAITTAACACAGAGCTCTCTCGGTCGCCACTTCACCTCAAG	1020
QY	1021	GCAAAAGGAGTCACTTGTACGTCGCGAGGAAGAGCCTTCTGATGGAA	1068
DB	1021	GCAAAAGGAGTCACTTGTACGTCGCGAGGAAGAGCCTTCTGATGGAA	1068
RESULT 3			
ID	AAC67965	standard; CDNA; 1149 BP.	
XX	AAC67965;		
DT	19-FEB-2001	(first entry)	
XX	Rat hepatoma H35 cell alpha1-2fucosyltransferase cDNA.		
KW	Rat; alpha1-2fucosyltransferase; cytostatic; neuroprotective;		
KW	neotropic; gene therapy; Fucalalpha1-2galbeta1-3GalNAc; immunotherapy;		
KW	immunosuppression; cancer; neurological disease;		
XX	small cell lung carcinoma; ss.		
OS	Rattus norvegicus.		
PN	WO200064464-A1.		
XX	02-NOV-2000.		
XX	23-APR-1999; 99WO-US07384.		
XX	23-APR-1999; 99WO-US07384.		
PA	(PACI-) PACIFIC NORTHWEST CANCER FOUND.		
PI	Holmes EH, Sherwood AL;		
DR	WPI; 2000-687252/67.		
DR	P-PSDB; AAB36104.		
XX	New rat ganglioside GM1-specific alpha1-2fucosyltransferase, useful for		
PT	preparation of fucosyl GM1 which is useful as a nutritional composition		
PT	or immunotherapeutic for cancer and neurological diseases -		
XX	Claim 10; Fig 5; 91pp; English.		
CC	The present sequence is given in a specification relating to a rat		
CC	ganglioside GM1-specific alpha1-2fucosyltransferase protein. The protein		
CC	or its cellular fraction is useful for synthesis of a molecule comprising		
CC	Fucalalpha1-2galbeta1-3GalNAc, a glycolipid, glycoprotein, glycolipoprotein		
CC	or a free oligosaccharide comprising Fucalalpha1-2galbeta1-3GalNAc.		
CC	The method involves contacting alpha1-2fucosyltransferase with GDP-fucose		
CC	and a molecule or glycolipid, glycoprotein, glycolipoprotein or		
CC	oligosaccharide having a terminal galbeta1-3GalNAc group. It is also		
CC	useful for synthesis of fucosyl-GM1 by contacting the protein with		

CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
CC glycolipoproteins, glycolipids and oligosaccharides are useful as
CC nutritional compositions and fucosyl-GM1 is useful for inducing an
CC immunotherapeutic or immunosuppressive action against cancer,
CC neurological disease or small cell lung carcinoma.

XX
SQ Sequence 1149 BP; 247 A; 341 C; 311 G; 250 T; 0 other;

Query Match 100.0%; Score 1068; DB 21; Length 1149;
Best Local Similarity 100.0%; Pred. No. 8.9e-289;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGACGGAATAGTGAAGCTCAACCCCTGTGACAGAGGAATACCGATGACGACT 60
DB 82 CTCACGACGGAATAGTGAAGCTCAACCCCTGTGACAGAGGAATACCGATGACGACT 141
QY 61 CAAATGTCTCGGGAACACAGAAAGCCAGAGATCGACGGACGACGACGATGGG 120
DB 142 CAAATGTCTCGGGAACACAGAAAGCCAGAGATCGACGGACGACGACGATGGG 201
QY 121 AATGGAGAGCTCGGGGCGATGTTACGATCAATTCATTGGCCGCTGGGAACGAGATG 180
DB 202 AATGGAGAGCTCGGGGCGATGTTACGATCAATTCATTGGCCGCTGGGAACGAGATG 261
QY 181 GCGGAATACGCCACACTCTTTGCACTGCGCAGGATGAACGACGCTTGCCTTCATCCCC 240
DB 262 GCGGAATACGCCACACTCTTTGCACTGCGCAGGATGAACGACGCTTGCCTTCATCCCC 321
QY 241 GCATCATGACACAGCTCTAGCCGCCATCTCAGGATCAGGCTCCCGTGTGTACACAGC 300
DB 322 GCATCATGACACAGCTCTAGCCGCCATCTCAGGATCAGGCTCCCGTGTGTACACAGC 381
QY 301 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 360
DB 382 GACACGGCCAAAAGATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGCGT 441
QY 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGGATACCCGTGCTCCGTGACCTTC 420
DB 442 TACCGCCACATTCGGGACACTTTGTGCGCTTCACGGGATACCCGTGCTCCGTGACCTTC 501
QY 421 TACCACCACTTCGGCCAGAGATCTGAGGAGTTCACCCCTGCATGACACGCTGGGGAG 480
DB 502 TACCACCACTTCGGCCAGAGATCTGAGGAGTTCACCCCTGCATGACACGCTGGGGAG 561
QY 481 GAGGCCACAGCCCTTCTCGTGTGCTGGGGTGAATGGGAGACCGACCGAGTACTTTTGTG 540
DB 562 GAGGCCACAGCCCTTCTCGTGTGCTGGGGTGAATGGGAGACCGACCGAGTACTTTTGTG 621
QY 541 GGTGTCCATGTGCGCCGAGGAGCTATGTGCATGTCTATGCTTAATGTGTGGAAGGCGTG 600
DB 622 GGTGTCCATGTGCGCCGAGGAGCTATGTGCATGTCTATGCTTAATGTGTGGAAGGCGTG 681
QY 601 GTGGCTGACGGGGTACTCGGAAAGCCCTGGATATGTTCCGGGACGCTATTATCTCT 660
DB 682 GTGGCTGACGGGGTACTCGGAAAGCCCTGGATATGTTCCGGGACGCTATTATCTCT 741
QY 661 CCAGTCTTCTGTTTACAGCAACCGGTATGGCTGTGCGGGAGAACATTAATGCTTCC 720
DB 742 CCAGTCTTCTGTTTACAGCAACCGGTATGGCTGTGCGGGAGAACATTAATGCTTCC 801
QY 721 CGAGGAGAGTGGTGTTCGGGGCAATGATGTAGGGGTGCCAGCCAGACGACTTCGGG 780
DB 802 CGAGGAGAGTGGTGTTCGGGGCAATGATGTAGGGGTGCCAGCCAGACGACTTCGGG 861
QY 781 CTGCTCACCCTGCAACACACACCATCATGACTATTGGACCTTGGGATTTGGGCTGCC 840
DB 862 CTGCTCACCCTGCAACACACACCATCATGACTATTGGACCTTGGGATTTGGGCTGCC 921
QY 841 TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCTTCCCGGATTCCTCGGTC 900
DB 922 TACCTGGCAGGTGGTATACCATCTACTTAGCCAACTACACCTTCCCGGATTCCTCGGTC 981
QY 901 CTCAAAGTCTTTAAGCCAGAGACGCTTCTTACCCGAAATGGGTGGGCATCCCTCCGAT 960

DB 982 CTCAAAGTCTTTAAGCCAGAGACCCCTTCTTACCCGAATGGTGGCATCCCTCCGAT 1041
QY 961 CTGTCCCACTCCCTTAAGGCATTACACACGAGCTGTCTCGGTCCCACTTCCACCTCAAG 1020
DB 1042 CTGTCCCACTCCCTTAAGGCATTACACACGAGCTGTCTCGGTCCCACTTCCACCTCAAG 1101
QY 1021 GCAAAAGGAGTCACTTGTAGTTCGACGAGAGAGCGCTTCTGATGGAA 1068
DB 1102 GCAAAAGGAGTCACTTGTAGTTCGACGAGAGAGCGCTTCTGATGGAA 1149

RESULT 4
AAD27207

ID AAD27207 standard; DNA; 1149 BP.

XX AAD27207;

XX 09-APR-2002 (first entry)

XX Rat hepatoma H35 cell alpha1-2Fuct DNA.

XX Rat: alpha1-2fucosyltransferase; alpha1-2Fuct; antisense therapy;
KW galactose beta1-3N-acetylglactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
KW cell transformation; ds.

XX Rattus norvegicus.

XX Key Location/Qualifiers

FT CDS 1..1143

FT /*tag= a

FT /product= "Rat hepatoma H35 cell alpha1-2Fuct"

FT misc_feature 662..1149

FT /*tag= b

FT /note= "Region which overlaps rat FTB"

XX US6329170-B1.

XX 11-DEC-2001.

XX 23-APR-1999; 990S-0298886.

XX 23-APR-1999; 990S-0298886.

XX (NWHO-) NORTHWEST HOSPITAL.

XX Holmes EH, Sherwood AL;

XX WPI; 2002-121132/16.

XX P-PSDB; AAE16622.

XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
PT GM1-specific alpha1-2fucosyltransferase enzyme by recombinant
PT techniques and for detecting oncogenic transformation of test tissues -

XX Claim 1; Fig 5; 41pp; English.

XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase
CC (alpha1-2Fuct) enzyme and its corresponding nucleic acid. This nucleic
CC acid is specific for a carbohydrate moiety found in ganglioside GM1,
CC a terminal galactose beta1-3N-acetylglactosamine (Galbeta1-3GalNAc)
CC saccharide. Alpha1-2Fuct DNA is useful for producing rat alpha1-2Fuct
CC protein by recombinant techniques. Alpha1-2Fuct DNA is useful for the
CC preparative synthesis of fucosyl containing glycolipids, glycoproteins,
CC glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.
CC Alpha1-2Fuct DNA is useful for detecting oncogenic transformation which
CC involves assaying for changes in expression of alpha1-2 Fuct. Since
CC alpha1-2Fuct is activated in cell transformation, antisense sequences
CC derived from alpha1-2Fuct DNA are useful for inhibiting, suppressing
CC or treating cancer. Alpha1-2Fuct DNA is useful in gene therapy and
CC antisense therapy. The present sequence is rat hepatoma H35 cell

CC	alpha1-2Fuct DNA.	
XX		
SQ	Sequence 1149 BP; 247 A; 341 C; 311 G; 250 T; 0 other;	
	Query Match 100.0%; Score 1068; DB 24; Length 1149;	
	Best Local Similarity 100.0%; Pred. No. 8.9e-289;	
	Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CTCGAGCAGCAATAGTGAAGCTCCAAACCCCTGTGCAGAGAAGGAATTACCGATGACGACT 60	
DB	82 CTCGAGCAGCAATAGTGAAGCTCCAAACCCCTGTGCAGAGAAGGAATTACCGATGACGACT 141	
QY	61 CAAATGTCTCTCGGAAACACAGAAACCCAGAGATCGACGGGACAGGAGCAGCATGGG 120	
DB	142 CAAATGTCTCTCGGAAACACAGAAACCCAGAGATCGACGGGACAGGAGCAGCATGGG 201	
QY	121 AATGGAGAGCTCGGGGGAATGTTACGATCAATTCATNTGGCCGCTGGGAAACAGATG 180	
DB	202 AATGGAGAGCTCGGGGGAATGTTACGATCAATTCATNTGGCCGCTGGGAAACAGATG 261	
QY	181 GCGAATACGCCACACTCTTTGCACTGGCCAGGATGAACGGACGGCTTGCCTTCATCCCC 240	
DB	262 GCGAATACGCCACACTCTTTGCACTGGCCAGGATGAACGGACGGCTTGCCTTCATCCCC 321	
QY	241 GCATCCATGCACACGCTCTAGCGCCCATCTTTCAGGATCAGCCTCCCGGTGTTACACAGC 300	
DB	322 GCATCCATGCACACGCTCTAGCGCCCATCTTTCAGGATCAGCCTCCCGGTGTTACACAGC 381	
QY	301 GACAGGGCAAAAGATCCCATGGCAGATTTACCATCTCAACGACTGGATGGAGAGCGT 360	
DB	382 GACAGGGCAAAAGATCCCATGGCAGATTTACCATCTCAACGACTGGATGGAGAGCGT 441	
QY	361 TACCGGCACATTCGGGACACTTTGTGCCTTCACGGGATACCGTGCCTCGACCTTC 420	
DB	442 TACCGGCACATTCGGGACACTTTGTGCCTTCACGGGATACCGTGCCTCGACCTTC 501	
QY	421 TACCACCACTCGCCCGACAGATCTCTGAAGAGTTTCAACCTCATGACACACGTGGGGAG 480	
DB	502 TACCACCACTCGCCCGACAGATCTCTGAAGAGTTTCAACCTCATGACACACGTGGGGAG 561	
QY	481 GAGGCCAGGCTTCTTGGTGTCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG 540	
DB	562 GAGGCCAGGCTTCTTGGTGTCTGGGGTGAATGGGAGCCAGCCGAGTACTTTTGTG 621	
QY	541 GTGTCTCAGTTCGCGGAGGAGACTATGTGCATGTATCCCTAATGTGGGAAGGCGGTG 600	
DB	622 GTGTCTCAGTTCGCGGAGGAGACTATGTGCATGTATCCCTAATGTGGGAAGGCGGTG 681	
QY	601 GTGGCTGACCGGGGTTACCTGGAAAAGCCCTGGATATGTTCCGGCAGCGCTATTCACT 660	
DB	682 GTGGCTGACCGGGGTTACCTGGAAAAGCCCTGGATATGTTCCGGCAGCGCTATTCACT 741	
QY	661 CCAGTCTTCTGTTTACAGCACGATATGGCTGTGTCGGGAGAACATTAATGTTCC 720	
DB	742 CCAGTCTTCTGTTTACAGCACGATATGGCTGTGTCGGGAGAACATTAATGTTCC 801	
QY	721 CGAGGAGAGCTGTGTTCGGGCAATGTTATGAGGGTGCAGCAGCAAGGACTTCGGG 780	
DB	802 CGAGGAGAGCTGTGTTCGGGCAATGTTATGAGGGTGCAGCAGCAAGGACTTCGGG 861	
QY	781 CTGCTCACCAGTGCACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 840	
DB	862 CTGCTCACCAGTGCACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCC 921	
QY	841 TACCTGTCAGGTGGTATACCATCTACTTAGCAACTACACCTTCGGGATTTCCGTTTC 900	
DB	922 TACCTGTCAGGTGGTATACCATCTACTTAGCAACTACACCTTCGGGATTTCCGTTTC 981	
QY	901 CTCAAAGTCTTTAAGCCAGGAGCAGCTTCTTCCCAAGTGGGCTGGCATCCCTGGCAT 960	
DB	982 CTCAAAGTCTTTAAGCCAGGAGCAGCTTCTTCCCAAGTGGGCTGGCATCCCTGGCAT 1041	
QY	961 CTGTGCCCACTCTTAAAGGCATTAACACCAAGCCTGTCTCGGTCCCACTTCCACCTCAAG 1020	

DB	1042 CTGTCCCACTCCTTAAGCATTAAACACAGCCTGTCTCGGTCCCACTTCCACCTCAAG 1101	
QY	1021 GCAAAAGAGGAGTCACTTGTACGTCGAGGAGAGCCTTCTGATGGAA 1068	
DB	1102 GCAAAAGAGGAGTCACTTGTACGTCGAGGAGAGCCTTCTGATGGAA 1149	
	RESULT 5	
	AAV29003	
ID	AAV29003 standard; DNA; 1043 BP.	
XX	AC AAV29003;	
XX	28-AUG-1998 (first entry)	
DE	Porcine secretor transferase (FUT2) gene.	
XX		
KW	Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;	
KW	transgenic animal; xenotransplantation; organ transplant; ss.	
XX		
OS	Sus scrofa.	
XX		
FH	Key Location/Qualifiers	
FT	9..1031	
FT	/*tag= a	
XX		
PN	WO9807837-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	22-AUG-1997; 97WO-AU00540.	
XX		
PR	23-AUG-1996; 96AU-0001823.	
XX		
PA	(AUST-) AUSTIN RES INST.	
XX		
PI	McKenzie IFC, Sandrin MS;	
XX		
DR	WPI; 1998-169148/15.	
XX	P-PSDB; AAM37855.	
PT	Nucleic acid encoding glycosyltransferase able to compete with	
PT	second such enzyme - particularly used to reduce expression of	
PT	unwanted carbohydrate epitope(s) on tissues intended for	
PT	transplantation	
XX		
PS	Claim 6; Fig 1A-B; 40pp; English.	
XX		
CC	This nucleotide sequence, designated clone pSe16.1, codes for	
CC	porcine secretor (Se) glycosyltransferase (see AAM37855), an enzyme	
CC	that has high affinity for type I and type III substrates. It was	
CC	isolated from a pig liver genomic library using full-length human	
CC	Se2 (FUT2) cDNA as probe. The nucleotide sequence shows about 36%	
CC	homology with human FUT1. Introduction of glycosyltransferase	
CC	nucleic acid, such as porcine Se nucleic acid, into a tissue	
CC	results in reduced expression of unwanted carbohydrate epitopes on	
CC	the tissue, especially porcine heart, liver, kidney or pancreas,	
CC	rendering it more suitable for transplantation, i.e. less	
CC	immunogenic and of increased immunological acceptability. A	
CC	claimed method of producing a cell from a donor species that is	
CC	immunologically acceptable to a recipient species involves reducing	
CC	levels of carbohydrate on the donor cell that causes it to be	
CC	recognised as non-self by the recipient by causing a nucleic acid	
CC	for a glycosyltransferase such as Se to be expressed in the cell.	
CC	Expression units, such as retroviral packaging or producer cells,	
CC	containing Se nucleic acids can be used in gene therapy.	
XX		
SQ	Sequence 1043 BP; 189 A; 352 C; 306 G; 196 T; 0 other;	
	Query Match 57.4%; Score 612.8; DB 19; Length 1043;	
	Best Local Similarity 80.6%; Pred. No. 2.3e-161;	
	Matches 716; Conservative 0; Mismatches 172; Indels 0; Gaps 0;	

QY	451	CAGTTACCCCTCATGACACAGCTGGGAGGAGCCAGGCCCTTCTGGTGGTCTGGG	510
Db	496	GAGTTACCCCTCATGACACAGCTGGGAGGAGCCAGGATTTCTGGGGTCTGGG	557
QY	511	GTGAATGGAGCAGCCAGTACTTTTGGTGTGCTCATGTGCGCGGAGGACATATG	570
Db	558	GTGAACGGAGCCGACCGAGTACCTACGTGGGGTGCACGTGCGCGGGGACATG	617
QY	571	CATGTCATGCCCTAAATGTGTGGAAGGCGTGTGCTACCGGGGTACCTGGAAGGCC	530
Db	618	CACGTGATGCCCAAGCTGTGGAAGGCGTGTGCGCGGACCGCGGTACCTGGAGCGCC	677
QY	631	CTGGATATGTTCCGGCAGCTATTACCTCCAGTCTTCGTTGTTTACAAGCAACGGTATG	690
Db	678	CTGGACTGTTTCCGGCTCGCTACCGCTCCCGCTTTGTGTCTCCAGCAACGGCATG	737
QY	691	GCCTGTGCGCGGAGAACATTAATGTTTCCCGAGGAGAGTGTGTCTCCGGCAATGTT	750
Db	738	GCCTGTGTGCGGAAACATCAATGCTTCGCGCGCGATGTGTGTTTCCCGCAATGCG	797
QY	751	ATGAGGGGTGCGCAGCAAGACATTCGGCTGCTCACCAGTGCACACACACCATCATG	810
Db	798	ATCGAGGGTCTCCCGCCAAAGACTTCGGGCTGCTCAGCAGTGTAAACACACTGTCTATG	857
QY	811	ACTATTGGGACCTTTTGGGATTTGGGCTGCTACCTGGCAGGTGGTGATACCATCTACTTA	870
Db	858	ACATTGGACAGTTCGGGATCTGGGCGCGCTACTTGTGTGGAGAGACCATCTACTTG	917
QY	871	GCCAACTACACCTTCCTGGATTCCTGGTTTCTCTCAAAAGTCTTTAAAGCAGAGCGCTTC	930
Db	918	GCCAAATTACAGCTCTCCGACTCTCCCTTCTCAAACTCTTTAAAGCCCGAGGAGCCTTC	977
QY	931	CTACCCGAATGGTGGGCTACCTGCGGATCTGCTCCGACTCTCCGACTCTTAAG	978
Db	978	CTCCCGAGTGGATTGGGATCGAGGAGAGCCTGTCCCGCACTCTCTTAAG	1025
RESULT 7			
AAV58323			
ID	AAV58323 standard; DNA; 2115 BP.		
XX	AAV58323;		
AC	AAV58323;		
XX	20-NOV-1998 (first entry)		
DT	Human Sec2 coding sequence.		
DE	Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;		
XX	GDP-L-fucose;beta-D-galactoside 2-alpha-L-fucosyltransferase; human;		
KW	FUT2; nonsecretor genotyping; ds.		
KW	Homo sapiens.		
XX	Key		
OS	Location/Qualifiers		
FH	64..1095		
FT	/*tag= a		
XX	US5807732-A.		
PN	15-SEP-1998.		
XX	28-FEB-1995; 95US-0395800.		
XX	28-FEB-1995; 95US-0395800.		
XX	(GIORGI) GIORGI D.		
PA	(KELL) KELLY R J.		
PA	(LENN) LENNON G.		
PA	(LOWE) LOWE J B.		
PA	(ROUQ) ROQUIER S.		
XX	Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;		

XX	WPI: 1998-520127/44.		
DR	P-PSDB; AAM69332.		
XX	DNA encoding fucosyltransferase enzyme - useful for producing recombinant enzyme and genotyping person as secretor or nonsecretor		
PT	Claim 1; Column 45-50; 55pp; English.		
PS	This sequence encodes the human Sec2 protein of the invention. The DNA encodes a alpha(1,2) fucosyltransferase and is the secretor.		
XX	H blood group alpha(1,2) fucosyltransferase gene. The DNA is useful for producing a recombinant human GDP-L-fucose:beta-D-galactoside		
CC	2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an individual as a secretor or nonsecretor as it is known that nonsecretors		
CC	homozygous for a mutant allele of the FUT2 gene that has a stop codon in the position corresponding to amino acid 143.		
XX	Sequence 2115 BP; 505 A; 606 C; 552 G; 452 T; 0 other;		
QY	128	AGCTGGCGGGCATGTTACAGATCAATTCATTTGGCGGCTGGGGACACAGATGGGGAAT	187
Db	239	AGTCTAGGGGATGTGGACGATCAATGCAATAGCCGCTGGGGAACACAGATGGCGAGT	298
QY	188	ACGCCACACATTCCTGGACCTGGCGAGGATGAACGAGCGGCTTGGTTCATCCCGCATCCA	247
Db	299	ACGCCACACTGTACGGCCTGGCCAAAGATGAACGGGGCGGCGCCCTTCATCCGCGCCAGA	358
QY	248	TGCACACAGCTCTAGCGCCCATCTTCAGGATCAGCCTCCCGGTGTTCACAGGAGACACGG	307
Db	359	TGCACAGACACCTTGGCGCCCATCTTCAGAAATCACTTGGCGGTGCTGCACAGCGCCACGG	418
QY	308	CCAAAAGATCCCATGGCAGATTAACCATCTCAACGACTGGATCGGAGGAGCGTTACCGCC	367
Db	419	CCAGCAGGATCCCTGGCAGAACTACCACCTGACGACTGGTGGAGGAGATACCGCC	478
QY	368	ACATTCGGGACACTTTTGTGCGCTTCACGGGATACCGGTCTCTCTGGAACCTTCTACCAAC	427
Db	479	ACATTCGGGGGAGTACGTCGCTTACCGGCTTACCGGCTTACCGCTCTCTGGAACCTTCTACCAAC	538
QY	428	ACCTGGCGCCAGAGATCTGAAGSAGTTCACTCCCTGATGATGACACGCTGCGGAGGAGGCC	487
Db	539	ACCTCCCGCAGGAGATCTCCAGGAGTTCACCTTGCACGACACGCTGCGGAGGAGGCC	598
QY	488	AGGCCTTCCTGCGTGTCTGCGGTGAATGGAGCCAGCCAGTACTTTTGTGGGTGCC	547
Db	599	AGAGTTCCTTGGGGGCTGCAGGTGAACGGAGCCGCGCGGACCTTTGTAGGGGTCC	658
QY	548	ATGTGCGCGAGGGAGTATGTGATGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG	607
Db	659	ATGTGCGCGAGGGAGTATGTGATGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG	718
QY	608	ACCGGGTTACCTGGGAAAGGCCCTGATATGTTCCGGGACACGCTATTCATCTCCACTCT	667
Db	719	ACCGGGATACCTACAGCAGGCGCTGGACTGGTTCGAGCTCGCTACAGCTCCCTCATCT	778
QY	668	TCGTGTGTACAAGCAACGGTATGGCCCTGGTGGCGGAGAACATTAATGCTTCCGAGGAG	727
Db	779	TCGTGTGTACAGTAATGGCATGSCCTGGTGTGCTGGGAGACATTTGACACCTCCACCGGTG	838
QY	728	ACGTGTGTTCGGGGCAATGGTATGAGGGGTGCGCCAGCAAGGACTTCCCGGTGCTCA	787
Db	839	ATGTGTGTTCGTGGCGATGGCATTCAGGGCTCACCTGGCAAAAGATTTTGTCTACTCA	898
QY	788	CCAGTGTCAACACACCATCATGACTATTTGGGACCTTTGGGATTTGGGCTGCTTACCTG	847
Db	899	CACAGTGTAAACACACCATCATGACTATTTGGGAGTTCGGGATTCGGGCGCCATACCTCA	958

	27-SEP-2000;	2000US-235840P.	
R	27-SEP-2000;	2000US-235863P.	
R	28-SEP-2000;	2000US-236028P.	
R	28-SEP-2000;	2000US-236032P.	
R	28-SEP-2000;	2000US-236033P.	
R	28-SEP-2000;	2000US-236034P.	
R	28-SEP-2000;	2000US-236109P.	
R	28-SEP-2000;	2000US-236111P.	
R	29-SEP-2000;	2000US-236842P.	
R	29-SEP-2000;	2000US-236891P.	
R	02-OCT-2000;	2000US-237172P.	
R	02-OCT-2000;	2000US-237173P.	
R	02-OCT-2000;	2000US-237278P.	
R	02-OCT-2000;	2000US-237294P.	
R	02-OCT-2000;	2000US-237295P.	
R	02-OCT-2000;	2000US-237316P.	
R	03-OCT-2000;	2000US-237425P.	
R	03-OCT-2000;	2000US-237598P.	
R	03-OCT-2000;	2000US-237604P.	
R	03-OCT-2000;	2000US-237606P.	
R	03-OCT-2000;	2000US-237608P.	
R	01-NOV-2000;	2000US-244867P.	
R	01-NOV-2000;	2000US-245084P.	
XX	(AVAL-) AVALON PHARM.		
PA	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horriqan S;		
XX	Soppet DR, Weaver Z;		
PFI	WPI; 2002-188264/24.		
DR			
XX	Screening for anti-neoplastic agent involves exposing cells to a		
XX	chemical agent to be tested for anti-neoplastic activity, and		
PFI	determining a change in expression of a gene of a signature gene set -		
XX	Claim 1; SEQ ID 4648; 44pp; English.		
XX	The present invention describes a method (M1) for screening for an		
CC	anti-neoplastic agent. The method involves exposing cells to a chemical		
CC	agent to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I) of a signature gene set, where (I)		
CC	comprises a sequence (S) selected from 8447 sequences (given in ABI61664		
CC	to AB170114), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening		
CC	an anti-neoplastic agent, and can be used for producing a product which		
CC	is the data collected with respect to the anti-neoplastic agent as a		
CC	result of M1, and the data is sufficient to convey the chemical		
CC	structure and/or properties of the agent. M1 can be used in the		
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,		
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer,		
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,		
CC	infiltrating lobular cancer, squamous cell carcinomas, neuroendocrine		
CC	carcinoma, papillary carcinoma and Wilm's tumour.		
XX	Sequence 3088 BP; 723 A; 839 C; 798 G; 728 T; 0 other;		
XX			
SQL			
	Query Match	55.2%; Score 589.4; DB 24; Length 3088;	
	Best Local Similarity	78.7%; Pred. No. 1.2e-154;	
	Matches 704; Conservative	0; Mismatches 191; Indels	0; Gaps
QY	128 AGCTCGGGGATGTTTCACGATCAATTCCATTGCCCGCTGGGAACCCAGATGGCGCAAT	187	
DB	287 AGCTCAGGGGATGTGGACGATCAATGCAATAGCGGCCTGGGGAACCCAGATGGCGCAAT	346	
QY	188 AGGCCACACTCTTTCACCTGCCAGGATGAACGACGGCTTGCGTTCATCCCGCATCCA	247	
DB	347 AGCCGACACTGTACGCCCTGGCCAAAGTAGACGGCGCGCGCTTCATCCCGCCCCAGA	406	
QY	248 TGSCAACAGGCTCTAGCGGCCATCTTCAGATCACAGCTCCCGGTTTACACAGCGCACGG	307	
DB	407 TGSCAGACACTGGCCCCCATCTTCAGATCACCTGCGGTTGCTGCAACAGCGCACGG	466	


```
SQ Sequence 3088 BP; 723 A; 839 C; 798 G; 728 T; 0 other;
Query Match 55.2%; Score 589.4; DB 24; Length 3088;
Best Local Similarity 78.7%; Pred. No. 1.2e-154;
Matches 704; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
QY 128 AGCTGGGGGATCTTCACCATCAATTCATTTGCCGGCTGGGAACACAGATGGCGAAT 187
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 287 AGCTAGGGGATGTGGACGATCAATCAATAGCCCGCTGGGAACACAGATGGCGAGT 346
QY 188 ACGCCACACTTTTGGCACTGGCCAGAGTAAGCGGCTTGGCTTCATCCCGGCATCCA 247
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 347 ACGCCACACTGTAGCCCTGGCCAAAGATGAAGGGCGCCGCTTCATCCGCGCCAGA 406
QY 248 TGCACACGCTCTAGGCCCACTCTTCCAGGATCAAGCTCCCGGGTGTACACAGGCACAGG 307
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 407 TGCACACAGCCCTGGCCCGCATCTTCAGAAATCAACCTTCCCGGTGCTCACAGGCCACGG 466
QY 308 CCAAAAAGATCCCATGGCAGAAATACCATCTCAAGCACTGATGGAGAGCGTTACCGCC 367
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 467 CCAGCAGATCCCTGGCAGACTACCACTCTGAAGCTGTGAGGAGGATATACCGCC 526
QY 368 ACATTCGGGACACTTTGTGGCTTACGGGATACCGCTCTCTGTGACCTTCTACACC 427
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 527 ACTTCCGGGGAGTACGTCGGTTCACCGCTACCGCTCTCTGTGACCTTCTACACC 586
QY 428 ACCTGGCCCGAGATCTCTGAAGAGTTTCACTCTGATGACACAGCTGGCGGAGAGGCC 487
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 587 ACCTCGCCCGAGAGATCTCTCCAGGAGTTTCACTCTGACGACCGCTGACGACCGCTGCGGGAGAGGCC 646
QY 488 AGCCCTCTCTCGCTGCTCTCGGGTGAATGGGACGACGAGTACTTTTGTGGGTGTC 547
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 547 AGAATTCCTTCGGGGCTGCAGGTGAACGGGACCGCGCGGACCTTTGTAGGGTCC 706
QY 548 ATGTGCGCGAGGGGACTATGTGCATGTATGCTTAATGTGTGAAGGGGCTGTGGCTG 607
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 707 ATGTGCGCGAGGGACTATGTCCATGTCTATGCCAAAAGTGTGGAAGGGGCTGTGGCG 766
QY 608 ACCGGGTATCCGGAAGAGCCCTGGATATGTTCGGGACACCTATCTCTCAGTCT 667
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 767 ACCGGGATACCTACAGAGCCCTGGATGTTCGGAGCTCGCTACAGCTCCCTCATCT 826
QY 668 TCCTGTATCAAGCAACGGTATGCTGCTGTCGGGAGACATTAATGCTCCCGAGGAG 727
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 827 TCCTGTACCACTAATGATGCTGCTGCTGTCGGGAGACATGACACCTCCACGGTG 886
QY 728 ACCTGTGTTCGGGGCAATGTATGAGGGTTCGGCGAGCAAGGACTTGGCGCTGCTCA 787
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 887 ATGTGTGTGTGCTGGCGATGGCATTGAGGGCTACCTGCCAAAGATTTTGTCTACTCA 946
QY 788 CCCAGTGCACACACACATCATGACTATTGGGACCTTTGGGATTTGGCTGCTACTGG 847
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 947 CACAGTGTACACACACATCATGACCTTTGGGACGTTTCGGGATCTGGCGGCATACCTCA 1006
QY 848 CAGGTGTGATACCACTACTTACTAGCAACTACACCTTCCGGATTTCCGTTCTCTCAAG 907
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1007 CGGGCGGAGACACATCTACCTGGCCANTTACACCTTCCCGACTCCCTTCTCTCAAA 1066
QY 908 TCTTTAAGCCAGAGGAGCGCTTCTTACCGAATGGGTGGGCATCTCCGTGCGAFTGTCC 967
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1067 TCTTTAAGCCAGAGGAGCGCTTCTTCCCGAGTGGACAGGATTTGGCGAGACCTGTCC 1126
QY 968 CACTCTCTTAGGCATTAACACACAGCTGCTGCTGGTCCACCTTCCACCTCAAGC 1022
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1127 CTTTACTCAAGCACTAATGCTGGCCCACTTTCTTGAGACCTTTTCTCTCTCTGTC 1181
```

RESULT 11

AAV21640

ID AAV21640 standard; cDNA; 1098 BP.

XX AC

XX AAV21640;

```
DT 08-JUL-1998 (first entry)
XX Plg H transferase encoding cDNA.
XX Plg; H transferase; chimeric; glycosyltransferase; gene therapy;
XX transplanted; ss.
XX Sus scrofa.
XX Location/Qualifiers
XX key 1.1098
XX CDS /*tag= a
XX /product= "H transferase"
XX WO9805768-A1.
XX 12-FEB-1998.
XX 01-AUG-1997; 97WO-AU00492.
XX 21-AUG-1996; 96US-0024279.
XX 02-AUG-1996; 96AU-0001402.
XX (AUST-) AUSTIN RES INST.
XX McKenzie IFC, Sandrin MS;
XX WPI; 1998-159170/14.
XX P-PSDB; AAW53102.
XX Nucleic acids encoding chimeric glycosyltransferases - used for
XX altering carbohydrate levels on the surface of cells, useful in gene
XX therapy and transplantation
XX Example 3; Fig 7; 51pp; English.
XX The present sequence encodes pig H transferase used in an example of the
XX present invention. The present invention describes nucleic acids (NA)
XX encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
XX domain of a first glycosyltransferase (GT) and a localisation signal of
XX a second GT, whereby when the NA is expressed in a cell and where the
XX chimeric enzyme is located in an area of the cell where it is able to
XX compete for substrate with a second GT, resulting in reduced levels of
XX a product from the second GT. The NAs can be used to produce cells and
XX organs with desired glycosylation patterns. Products and methods of the
XX present invention can be used to reduce the levels of undesirable
XX epitopes in cells, tissues or organs which may be used in
XX transplantation or gene therapy.
XX Sequence 1098 BP; 181 A; 367 C; 312 G; 238 T; 0 other;
```

```
Query Match 39.7%; Score 423.8; DB 19; Length 1098;
Best Local Similarity 59.6%; Pred. No. 1.9e-108;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;
QY 136 GGCAATGTTACAGATCAATTCATTTGCCGGCTGGGAACACAGATGGCGAATACGCCACA 195
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 235 GGGACCTGGACTATTTACCGGATGGCGGTTTGGGAACACAGATGGCGAATACGCCACG 294
QY 196 CTCTTTTCACTGGCAGGATGAACGGACGGCTTTCATCCCGCATCCCATCCATGCACAC 255
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 295 CTCTGGCTTGGCGAGCTCAACGGCGCCAGCGCTTCATCCAGCTTGCATGCGACGCC 354
QY 256 GCTTACGCCCCATCTTCAGGATCAGCGCTCCCGGTGTTCACACAGCAGCAGCGCAAAAG 315
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 355 GTCTTGGCCCCCGTGTTCGCAATCAGCTGCTGTCTCTGGCGCCGAGGTAGACAGGCAC 414
QY 316 ATCCCATGGCAGAAATTAACATCTCAACGACTGGATGAGAGCGGTTACCGCCACATTCGG 375
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 415 GCTCCTTGGCGGAGCTGGAGCTTACGACTGGATGCTCGAGGATTATGCCCACTTAAAG 474
QY 376 GGACACTTGTGGCTTTCACGGGATACCGGTGCTCTCTGGACCTTTCACACACCTTGGCG 435
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 475 GAGCCTGGCTGAAGCTCACGGCTTCCCTGCTCTGAGCTTCTTCCACCAACCTCCGG 534
QY 436 CCAGAGATCCCTGAAGAGCTTACCTGTCATGACCACTGCGGAGGAGGCCAGCCAGCCCTTC 495
Db 535 GAGCAGATCCCGAGGAGTTCACCTGTCACGACCTTTCGGCAAGAGGCCAGGGGTA 594
QY 496 CTGCGT-----GGTCTCGGGTGAATGGAGGCCAGCCGAGTACTTTTGTGGGTCTCCAT 549
Db 595 CTGAGTCAGTTCCTGCTACCCCGCACAGGGACCCGCCAGACACTTCTGGGGTCCAC 654
QY 550 GTGCGCCGAGGGGACTATGTGATGTCATGCTAATGTGTGGAAGGGCGTGGTGGCTGAC 609
Db 655 GTGCGCCGCGGAGTATCTGCTGTGATGCCAAGACGCTGGAAGGGGTGTGGGTGAC 714
QY 610 CGGGTTTACCTGGAAAGCCCTGGATATGTTCCGGGACGCTATTCATCTCCAGTCTTC 669
Db 715 GCGCTTACCTCCAGCAGGCTATGACATGTTCCGGGCCGATGACGAAGCCCGCTTT 774
QY 670 GTGGTTACAAGCAACGGTATGCGCTGTGCGGGAGAACATTAATGCTTCCGAGGAGAC 729
Db 775 GTGGTCAACCAAGCAACGGTATGAGTGTGCGGAAGAACATCGACACCTCCCGGGGGAC 834
QY 730 GTGGTGTTCGGGGCAAGTATGAGGGTGTGCGGAGGCTTCCGCTGCTCACC 789
Db 835 GTGATCTTTGTGGGATGGGGGAGGCGCGCCCGCAGGACTTTGCGCTGCTGGTG 894
QY 790 CAGTGCACCAACCAACCATATGACTATTGGGACCTTTGGGATTTGGCTTCCCTACCTGGCA 849
Db 895 CAGTGCACCAACCAACCATATGACCAATGGCACCCTTCGGCTTCGGCCGCTACCTGGCT 954
QY 850 GTGGTGTATACCATCTACTAGCAACATACACCTTCCGGATTCCTCGTTCCTCAAGTC 909
Db 955 GTGGAGATACCATCTACTTGGCTTAACCTACCTTCCCTTCCGCTTCCCTGAAGTC 1014
QY 910 TTAAAGCAGAGGAGCCTTCTACCCGATGGGTGGGATCCCTGCGATCTGTCCTCA 969
Db 1015 TTAAACCGAGGCTTCTCTCGGAGTGGGTGGGCAATTAATGACAGACTTGTCTCA 1074
QY 970 CTCCTTAAG 978
Db 1075 CTCAGATG 1083

RESULT 12
ID AAX15872
AC AAX15872 standard; DNA; 1269 BP.
CC AAX15872;
XX AAX15872;
DT 13-MAY-1999 (first entry)
DE Fragment of pHT83xF encoding swine alpha(1,2) fucosyltransferase.
KW Swine alpha(1,2) fucosyltransferase; FUT1; Escherichia coli; resistant;
KW E. coli-associated intestinal disorder; E coli infection; ss.
XX
OS Synthetic.
OS Sus sp.
XX
FH Key
FT CDS
FT CDS
XX
XX W09853101-A2.
XX
XX 26-NOV-1998.
XX
XX 20-MAY-1998; 98WO-US10259.
XX
XX 20-MAY-1997; 97US-0047181.
XX
XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
```

```
PI Bosworth BT;
XX
DR WPI: 1999-131692/11.
DR P-PSDB; AAW97356.
XX
PT Identifying swine genetically resistant to E. coli associated
PT diseases - using PCR-RFLP to assay for polymorphisms in the
PT alpha(1,2) fucosyltransferase 1 gene
XX
PS Claim 6; Fig 1; 19pp; English.
XX
CC The present sequence encodes swine alpha(1,2) fucosyltransferase
CC (FUT1). The specification describes methods relating to Escherichia
CC coli-resistant swine. One of the methods for identifying a swine
CC resistant to E. coli-associated intestinal disorders, comprises
CC determining whether the base at 307 of alpha(1,2) fucosyltransferase
CC gene (FUT1) is adenine (sic), in which case the swine are resistant. The
CC porcine FUT1 polymorphisms can be used to develop drugs for the
CC treatment of swine having E. coli-associated disease. The methods can
CC also be used in breeding programmes to identify swine with resistance
CC to E coli infection.
XX
SQ Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;
Query Match 39.7%; Score 423.8; DB 20; Length 1269;
Best Local Similarity 69.6%; Pred. No. 2e-108;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;
QY 136 GCATGTTCAAGATCAATTCATTCGCGGCTGGGGAACCAAGATGGGGAATACGCCACA 195
Db 243 GGGAGCTGGAGTATTTACCGGATGGCGGTTTGGGAACCAAGATGGGAGTATGCCAG 302
QY 196 CTCCTTGGACTGCGCAGGATGAACGGGCTTGCATCCCGCATCCCGCATGCACAAAC 255
Db 303 CTGCTGGCCCTGGCGAGCTCAACGGCCCGCAGGCTTCATCCAGCTCCATGACGCC 362
QY 256 GCTCTAGCGCCCATCTTTCAGGATCAGCCTCCCGGTGTTACACAGGAGACAGCCAAAAG 315
Db 363 GTCCTGGCCCGCTGTTCCGATCAGCTGCTGCTGCTGGCCCGCGAGGTAGAGGCAC 422
QY 316 ATCCCATGGCAGAATACCATCTCAACGACTGGATGGAGGAGCGTTACGCCACATTCG 375
Db 423 GCTCCTTGGCGGAGCTGGAGCTTCACGACTGGATGTCGAGGATATGCCACTTAAG 482
QY 376 GGACACTTTGTGCGCTTTCACGGATACCGTCTGCTGACCTTCTACCAACACCTGCGC 435
Db 483 GAGCCCTGGCTGAAGCTACCGCTTCCCTGCTGAGCTTCTTCCACCACTCCCG 542
QY 436 CCAGAGATCCTGAAGAGTTCACCTGTCATGACCTTAATGTGTGGAAGGCGCTGGTGGT 495
Db 543 GAGCAGATCCCGAGCGAGTTCACCTGTCACGACCACTTCGGCAAGAGGCCCGGGTA 502
QY 496 CTGCGT-----GGTCTCGGGTGAATGGAGGCCAGCCAGTACTTTTGGGTGTCCAT 549
Db 603 CTGAGTCAAGTTCGCTTACCCGACAGGGGACCGCCACGACCTTCGTGGGGTCCAC 562
QY 550 GTGCGCCGAGGGGACTATGTGCTATGCTCAATGCTTAATGTGTGGAAGGCGCTGGTGGT 609
Db 663 GTGCGCCGCGGGACTATCTGCTGATGCCAGCGCTGGAGGGGTGGTGGGTGAC 722
QY 610 CGGGGTACCTGGAAGGCCCTGGATATGTTCCGGGACAGCTATTCATCTCAAGTCTTC 669
Db 723 GCGCGTTACCTCCAGCAGCTATGGTGTTCGGGCGCGGATAGCAAGCCCGCTCTT 782
QY 670 GTGGTTACAAGCAACGGTATGGCTTGGTCCCGGAGAGACATTAATGCTTCCCGAGGAC 729
Db 783 GTGGTCAACCAAGCAACGGTATGGAGTGGTCCCGGAGAGACATCGACCTTCCCGGGGAC 842
QY 730 GTGGTGTTCGGGGCAATGGTATTTGAGGGTGTCCAGCGCAAGGACTTCGCGTGTCCAC 789
Db 843 GTGATCTTTGCTGGCGATGGGCGGAGGCGCGCCCGCAGGACTTTCGGTGTGGTG 902
QY 790 CAGTGCACCAACCAACCATCATGACTATTGGGACCTTTGGGACTTTGGGCTGCTACCTGCA 849
```

[illegible]

RESULT 13

AA03811

ID AAX03811 standard; DNA; 1269 BP.

XX

XX AAX03811;

XX AC

XX DT

XX 01-APR-1999 (first entry)

XX

XX Swine alpha-1,2-fucosyltransferase 1 encoding DNA.

DE

XX

XX Swine; pig; alpha-1,2-fucosyltransferase 1; FUT1; resistance;

KW Escherichia coli; infection; oedema; postweaning diarrhoea;

KW

KW intestinal disorder; polymorphism; ss.

XX

XX Sus scrofa.

XX

XX

XX Key Location/Qualifiers

FH 9..1106

FT

FT /*tag= a

XX

XX W09853102-A1.

XX

XX

XX 26-NOV-1998.

PD

XX

XX

XX 20-MAY-1998; 98WO-US10318.

PF

XX

XX 20-MAY-1997; 97US-0047181.

PR

XX

XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.

PA

PA (SWTE-) SWISS FEDERAL INST TECHNOLOGY ZURICH.

PA (USDA) US SEC OF AGRIC.

XX

XX Bosworth BT, Vogeli P;

PI

XX

XX WPI: 1999-059746/05.

DR

DR P-PSDB; AAW30630.

XX

XX

XX New method of identifying swine that are resistant to intestinal

PT colonisation by Escherichia coli - comprises use of genetic

PT polymorphic markers, used for breeding swine resistant to

PT Escherichia coli-related diseases

XX

XX

XX Claim 6; Fig 1; 35pp; English.

PS

XX

XX A method has been developed for the identification of swine that are

CC resistant to intestinal colonisation by E. coli. The method comprises

CC determining whether a genetic polymorphism associated with resistance to

CC colonisation is present in a swine sample, and then inferring that the

CC swine is resistant if it is homozygous for the polymorphism. The method

CC uses the swine alpha-1,2-fucosyltransferase (FUT1) polymorphism. The

CC present sequence encodes swine FUT1. The method enables the breeding of

CC swine that are resistant to E. coli-related diseases. This method

CC comprises breeding swine that have a genetic polymorphism in the FUT1

CC gene. More particularly, the identification method identifies swine that

CC are resistant to E. coli-related intestinal disorders if, in a sample

CC taken, the only nitrogen base at residue 307 in the FUT1 gene is adenine.

CC Larger amplified fragments from the assay can be used for RFLP analysis,

```

Db      1083 CTCGAGATG 1091
|||||
RESULT 14
AAZ94417
ID AAZ94417 standard; DNA; 1269 BP.
XX
AC AAZ94417;
XX
XX
DT 18-JUL-2000 (first entry)
XX
DE Pig alpha-1-2 fucosyltransferase FUT1 gene.
XX
KW Alpha-1-2 fucosyltransferase; FUT1; pig; polymorphism;
KW Escherichia coli; resistance; ss.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT CDS 9..1106
FT /tag= a
FT /transl_except= (pos:726..728, aa:Ala)
FT variation replace(315,A)
FT /tag= b
FT /note= "adenine is substituted for guanine in
FT resistant pigs"
XX
XX WO200016641-A1.
XX
XX 30-MAR-2000.
XX
XX PF 17-SEP-1999; 99WO-US21408.
XX
XX 18-SEP-1998; 98US-0151592.
XX
XX (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
XX (USDA ) US SEC OF AGRIC.
XX
XX Bosworth PT, Ridpath J, Wiseman B;
XX WPI: 2000-283404/24.
XX DR P-PSDB; AAY79302.
XX
XX Improving weight gain in swine using swine genetically resistant
XX Escherichia coli and feeding swine high levels of plant based protein
XX
XX Disclosure; Fig 1; 33pp; English.
XX
XX This is the nucleotide sequence of the pig FUT1 gene encoding
XX alpha-1-2 fucosyltransferase (see AAY79302). A polymorphism at
XX position 307 of the coding region is associated with susceptibility
XX to F18 Escherichia coli colonization; pigs homozygous for adenine
XX at position 307 are resistant to colonization, while heterozygous
XX animals and animals homozygous for guanine at position 307 are
XX susceptible to colonization. A claimed method for improving weight
XX gain in pigs involves selecting animals that are genetically
XX resistant to E. coli colonization and feeding these animals high
XX levels of plant-based protein concentrate. A claimed method for
XX preventing F18 E. coli colonization in swine, especially swine that
XX are genetically susceptible to F18 E. coli colonization, involves
XX replacing some or all of the plant-based proteins in the diet with
XX animal-based proteins. The polymorphism in the FUT1 gene is also
XX useful for developing drugs to treat swine that have E. coli
XX associated disease. The polymorphism can be detected using
XX PCR-RFLP tests (see also AAZ94418-19).
XX
XX Sequence 1269 BP; 219 A; 413 C; 362 G; 275 T; 0 other;
XX
XX Query Match 39.7%; Score 423.8; DB 21; Length 1269;
XX Best Local Similarity 69.6%; Pred. No. 2e-108;
XX Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

```

```

QY 136 GGCATGTTCAAGATCAATTCCATTGGCCGGCTGGGAAACACAGATGGCGAATACGCCACA 195
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 GGGACCTGGACTATTATCCCGGATGGCCGGTTGGGAACACAGATGGAGATGCCACG 302
QY 196 CTCCTTGCACCTGCCAGGATGAACGACGGCTTGCCTTCATCCCCCATCCATGCACAAC 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 CTGCTGGCCCTGGCGCAGCTCAACGGCCGCCAGGCTTCATCCAGCTCCCATGCACGCC 362
QY 256 GCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTTACACAGGACACAGCCCAAAAG 315
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 363 GTCCCTGGCCCGCGTGTTCGGCATCAGCTGCTGCTGCGCCGCCGAGGTACAGGCAC 422
QY 316 ATCCCATGGCAGAATTACCATCTCAAGGACTGATGGAGGAGCGTTACCGCCACATCCG 375
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 423 GCTCCTTGGCGGAGCTGGAGCTCAGGACTGATGTCGGAGGATTATGCCCACTTAAG 482
QY 376 GGACACTTTGTGCGCTTCACGGGATACCGGTCTCTCTGGACCTTCTACCAACACCTGGCC 435
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 GAGCCCTGGCTGAAGCTCACCGCTTCCCTCTCTCTGGACCTTCTTCCACCACTCCGG 542
QY 436 CCAGAGATCCTGAAGGAGTTACCCCTGCATGACAGCTGCGGGAGAGCCCGCAGCCCTTC 495
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 543 GAGCAGATCCGACGCGAGTTACCCCTGCACGACCACTTCGGCAAGAGAGCCCGAGGGGTA 602
QY 496 CTGCGT-----GGTCTGCGGTGAATGGGAGCCAGCCGAGTACTTTTGTGGGTCTGCAT 549
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 CTGAGTCTAGTCCCTCTACCCCGCACAGGGNACCCGCCAGCACCTTCGTGGGGTCCAC 662
QY 550 GTGCCCGAGGGGACTATGTCATGTCATGCTTAATGTGGAAGGGGGTGGTGGCTGAC 609
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 663 GTGCCCGCGGGGACTATCTGCTGTGATGCCAAGCGCTGGAAGGGGGTGGTGGGAGAC 722
QY 610 CGGGTTACCTGGAAAAGCCCTGGATGATGTCGGGACAGCTATTCATTCACCTCTTC 669
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 723 GGCGTTTACCTCCAGCAGGCTATGGACTGGTTCGGGGCCCGATAGCAAGCCCGCTCTTT 782
QY 670 GTGTTACAAGCAACGGTATGGCCCTGGTCCGGGAGAACATTAAATGCTTCCGAGGAGAC 729
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 783 GTGTTACCAAGCAACGGTATGGCCCTGGTCCGGGAGAACATTAAATGCTTCCGAGGAGAC 842
QY 730 GTGGGTTTCGGGGCAATGGTATGAGGGGTGCGCAGCCAGGACTTCGGCTGCTCACC 789
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 843 GTGATCTTTTGTGGCGATGGCGGGAGCGCGCGCCCGCAGGGACTTTGCGCTGCTGGTG 902
QY 790 CAGTGCACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCGCTACCTGCA 849
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 903 CAGTGCACACACACCATCATGACTATTGGCCTTCGGCTTCTGGGCGCGCTACTCGCT 962
QY 850 GGTGTTGATACCATCTACTTAGCCCAACTACACCTTCGGGATTCCTCCCTCAAGATC 909
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 953 GGTGAGATACCATCTACTTAGCTAACTTACCTGCCCCACTTCACGCTTCTTGAAGATC 1022
QY 910 TTTAAGCCAGAGGAGCGCTTCCTACCCGAATGGGTGGGCAATCCCTGCGCATCTGTCCCA 969
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1023 TTTAAACCCGAGGTGCTCTCTGCCGATGGGTGGGCATTAAATGCAGACTTGTCTCCA 1082
QY 970 CTCCTTAAG 978
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1083 CTCGAGATG 1091

```

RESULT 15

AAI72831

ID AAI72831 standard; cDNA; 1269 BP.

XX AC AAI72831;

XX DT 22-JUL-2002 (first entry)

XX DE FUT1 cDNA.

XX KW Gene; pig; swine; alpha (1,2) fucosyltransferase; FUT1; weight gain;

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	1068	100.0	1068	4	US-09-298-886-9		Sequence 9, Appli
2	1068	100.0	1149	4	US-09-298-886-7		Sequence 7, Appli
3	612.8	57.4	1043	4	US-09-254-077A-5		Sequence 5, Appli
4	590.6	55.3	2115	1	US-08-395-800A-7		Sequence 7, Appli
5	516.6	48.4	1144	1	US-08-395-800A-1		Sequence 1, Appli
6	423.8	39.7	1269	1	US-09-151-593-1		Sequence 1, Appli
7	408.2	38.2	1155	1	US-08-208-889A-1		Sequence 1, Appli
8	408.2	38.2	1155	2	US-08-433-271-1		Sequence 1, Appli
9	405	37.9	999	5	PCR-US914-00899-10		Sequence 10, Appli
10	405	37.9	1136	1	US-08-395-800A-9		Sequence 9, Appli
11	405	37.9	1174	5	PCR-US959-07554-3		Sequence 3, Appli
12	405	37.9	1199	1	US-08-395-800A-5		Sequence 5, Appli
13	405	37.9	2268	4	US-08-675-773B-4		Sequence 4, Appli
14	405	37.9	3373	1	US-08-273-411-2		Sequence 2, Appli
15	405	37.9	3791	4	US-08-675-773B-3		Sequence 3, Appli
16	405	37.9	8174	1	US-07-914-281-5		Sequence 5, Appli
17	405	37.9	8174	1	US-08-393-246-5		Sequence 5, Appli
18	405	37.9	8174	1	US-08-525-058A-5		Sequence 5, Appli
19	405	37.9	8174	2	US-08-696-731-5		Sequence 5, Appli
20	405	37.9	8174	4	US-09-043-531-5		Sequence 5, Appli
21	405	37.9	8174	5	PCR-US914-00899-3		Sequence 3, Appli
22	403.4	37.8	1155	1	US-08-434-151-1		Sequence 1, Appli
23	403.4	37.8	1155	2	US-08-715-259-1		Sequence 1, Appli
24	72.8	6.8	100	4	US-09-298-886-29		Sequence 29, Appli
25	71.2	6.7	100	4	US-09-298-886-23		Sequence 23, Appli
26	71.2	6.7	100	4	US-09-298-886-26		Sequence 26, Appli
27	69.6	6.5	100	4	US-09-298-886-20		Sequence 20, Appli


```
QY 670 GTGTTACAGCAACGGTATGGCTGGTGGGAGAAATTAATGCTTCCCGAGGAGAC 729
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 GTGGTCAACGCGATGACATGGCTGGTGGGAGAGCATACAGCTCCCTTGGGGAC 839
QY 730 GTGGTGTCCGGGGAATGTAATGAGGGGTCCCGAGCCAAAGACTTCGGGCTGCTCAAC 789
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 840 GTGGTGTTCCTCGCAATGCGCTCCAGGGCTACCTGCCAAGACTTGGCACTGCTCACA 899
QY 790 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTCGCA 849
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 900 CAGTGAACACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTCGCG 959
QY 850 GGTGGTATACCATCTACTTAGCAACTACACCTTCCGGATTTCCGTTCCCTCAAAGTC 909
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 960 GGGGGGACACTGTCTACCTGGGCACTTCAACCTTCCCAACTCCCTTTCAAGTGGTC 1019
QY 910 TTAAACCCAGAGCGAGCTTCTTACCCGAATGGGTGGGCAATCCCTGCCGATCT 962
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 TTAAACCCAGAGCGAGCTTCTTACCCGAATGGGTGGGCAATCCCTGCCGATCT 1072

RESULT 6
US-09-151-592-1
; Sequence 1, Application US/09151592
; Patent No. 6355859
; GENERAL INFORMATION:
; APPLICANT: Bosworth, Brad
; APPLICANT: Ridpath, Julia
; APPLICANT: Wiseman, Barry
; TITLE OF INVENTION: INTERACTIONS BETWEEN GENOTYPE AND DIET IN SWINE THAT
; PREVENT E. COLI ASSOCIATED INTESTINAL DISEASE
; FILE REFERENCE: 21419-90119
; CURRENT APPLICATION NUMBER: US/09/151,592
; CURRENT FILING DATE: 1999-09-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Swine
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1103)
; FEATURE:
; OTHER INFORMATION: FUT1
US-09-151-592-1

Query Match 39.7%; Score 423.8; DB 4; Length 1269;
Best Local Similarity 69.6%; Pred. No. 2.2e-115;
Matches 591; Conservative 0; Mismatches 252; Indels 6; Gaps 1;

QY 136 GGCATGTTACAGTATCAATTCATTGCGCGCTGGGAACAGATGGCGATAGCCACA 195
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GGCATGAGTATTTACCGGATGGCGGTGGGAACAGATGGAGATAGCCACG 302
QY 196 CTCCTTTGCTGCGCCAGGATGAACGAGCGCTTGGCTTCATCCCGCATCCATGCAAC 255
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 CTCCTGGCCCTGGCGAGTCAACGCGCGCGAGCGCTTCATCCAGCTGCGATGCGCGC 362
QY 256 GCTCTAGGCGCCATCTCAGGATFACCGCTCCCGGTGTTACACAGCGACAGCGCAAAAG 315
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GTCCTGGCGCCCGTGTTCGCACTACGCTGCCGTCTCGCGCGCGAGGATGAGAGGCAC 422
QY 316 ATCCCATGGCAGATTTACCATCTCAGGACTGATGAGGAGCGTTACCGCCACATCCG 375
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GCTCCCTGGCGGAGTGGAGCTTCAGACTGATGATGCCAGGATTTGCCCACTTAAAG 482
QY 376 GGACACTTTTGGCGCTTACGGGATACCGGTGCTCTCGGACCTTTCACACCACTCGCG 435
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GAGCCCTGGCTGAGTCAACCGGCTTCCCTGCTGCTGACCTTCTTCCACCACTCCGG 542
QY 436 CCAGAGATCCTGAAGGATTCACCTTCATGACCACTGGGAGGAGGCCAGGCTTC 495
```

```
Db 543 GACAGATCGCAGGAGTTCACCTCGACGACACCTTCGGCAAGAGGCCCGGGGTA 602
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 CTGGT-----GGTTCGGGTGAATGGAGCCAGCCGAGTACTTTGTGGGTGTCAT 549
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 CTGAGTCAGTTCGCTTACCCCGCACAGGGGACCGCCCGACCTTTCGTGGGGTCCAC 662
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 550 GTCGCGCGAGGGACATGTCATGTCATGCTAAATGTGTGGAGGGCGTGGTGGCTGAC 609
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 GTCGCGCGCGGAGTATCTCGTGTGATCCCAAGCGCTGGAGGGGGTGGTGGGTGAC 722
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 610 CGGGGTTCACCTGGAAGAGCGCTTGGATATTTCCGGGCAAGCTATTATCTCCAGTCTTC 669
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 GCGCTTACCTCCAGCAGGCTAIGGACTGTTCCGGGCCGATAGGAAGCCCGCTCTIT 782
QY 670 GTGGTTACAGCAACGGTATGCGCTTGTGCGCGGAGAACATTAATGCTTCCAGAGGAC 729
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 GTGGTCAACAGCAACGCGATGGTGTGCGGAGAACATCGACACTCCCGGGGGAC 842
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 730 GTGGTGTTCGCGGCAATGATTAGGGGTGCGCAGCAAGGACTTTCGCGCTGCTCAC 789
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 GTGATCTTGTGGGATGGCGGAGCGCGCCCGCAGGACITTTGCGTCTGGTG 902
QY 790 CAGTGAACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTGGCA 849
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 CAGTGAACACACCATCATGACTATTGGGACCTTTGGGCTTCTGGGCGCGCTACCTGGCT 962
QY 850 GGTGGTATACCATCTACTTAGCAACTACACCTTCCGGATTTCCGGTCTCTCAAGTC 909
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GGTGGAGATACCATCTACTTGGGTAACCTTACCGCTGCCACTTCCAGCTTCCCTGAGATC 1022
QY 910 TTAAAGCCAGAGCGAGCTTCTTACCGGAATGGGTGGGCAATCCCTGCGGATCTGCCCCA 969
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 TTAAAGCCAGAGCGAGCTTCTTCCCGCGAGTGGTGGGCAATTAATGAGACTTGTCTCCA 1082
QY 970 CTCCTTAAAG 978
||||| ||| |||
Db 1083 CTCAGATG 1091

RESULT 7
US-08-208-889A-1
; Sequence 1, Application US/08208889A
; Patent No. 5750176
; GENERAL INFORMATION:
; APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,
; APPLICANT: Moremen, K., Pierce, J.
; TITLE OF INVENTION: Transgenic Production of Oligosaccharides and
; TITLE OF INVENTION: Glycoconjugates
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Donald O. Nickey
; ADDRESSEE: ROSS Products Division
; ADDRESSEE: Abbott Laboratories
; STREET: 625 Cleveland Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: United States of America
; ZIP: 43215
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS Version 6.21
; SOFTWARE: WordPerfect Version 6.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,889A
; FILING DATE: 09-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: No. 5750176 applicable
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614) 624-7080
; TELEFAX: (614) 624-3074
; TELEX: No. 5750176e
```

```

/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1155 base pairs
/ TYPE: Nucleic acid
/ STRANDEDNESS: Single
/ TOPOLOGY: Unknown
/ MOLECULE TYPE: Cloned cDNA representing the product of a
/ MOLECULE TYPE: human genomic DNA segment
/ DESCRIPTION: GDP-L-fucose- D-galactoside 2-alpha-fucosyl-transferase
/ HYPOTHEICAL:
/ ANTI-SENSE:
/ FRAGMENT TYPE: Entire amino acid sequence provided.
/ ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line
/ ORGANISM:
/ STRAIN:
/ INDIVIDUAL ISOLATE:
/ DEVELOPMENTAL STAGE:
/ HAPLOTYPE:
/ TISSUE TYPE:
/ CELL TYPE:
/ CELL LINE:
/ ORGANELLE:
/ IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line
/ LIBRARY:
/ CLONE:
/ POSITION IN GENOME:
/ CHROMOSOME/SEGMENT: 19
/ MAP POSITION:
/ UNITS:
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD: DNA sequencing and restriction analysis
/ OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:
/ OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- D-galactoside 2-alpha-fucosyl-
/ OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO:1:
/ OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.
/ PUBLICATION INFORMATION:
/ AUTHORS: V.P. Rajan, et al.
/ TITLE:
/ JOURNAL: Journal of Biological Chemistry
/ VOLUME: 264
/ ISSUE:
/ PAGES: 11158 - 11167
/ DATE: 1989
/ DOCUMENT NUMBER:
/ FILING DATE:
/ PUBLICATION DATE:
/ RELEVANT RESIDUES IN SEQ ID NO:
/
/ US-08-208-889A-1
/
/ Query Match 38.2%; Score 408.2; DB 1; Length 1155;
/ Best Local Similarity 68.7%; Pred. No. 8.5e-111;
/ Matches 579; Conservative 0; Mismatches 258; Indels 6; Gaps 1;
/
/ QY 136 GGCATGTCACGATCAATTCCTATGGCGGCTGGGGAACACGATGGGGGACGACAC 195
/ DB 272 GGCACCTGGACTGTCTACCCCAATGGCGGTTGGTAAATCAGATGGGACAGTATGCCAG 331
/ QY 196 CTCCTTGCACTGCCAGGATGACGACGCGCTGGCTTCATCCCGCATCCATGCACAC 255
/ DB 332 CTGCTGGCTCTGGCCAGCTCAACGCGCGCGGCGCTTATCTGCTGCCATGATGCC 391
/ QY 256 GCTCTAGCCGCAATCTCAGGATCAGCCCTCCCGGTGTACACAGGACACGCGCCAAAAG 315
/ DB 392 GCCCTGGCCCCGGTATTCGGATCACCCTCCCGTGTGGCCGCGGAGAGTGACACGCC 451
/ QY 316 ATCCCATGACGAAATFACATCTCAACGACTGGATGGAGGAGCGTTTACGCCACATTCGG 375
/ DB 452 ACGCCGTGGCGGAGTGTCAGCTICAGACTGGATGCTCGGAGGAGTACGCGACTTGAGA 511
/ QY 376 GGACACTTTGCGGCTTACGGGATACCGGTGCTGCTGACCGCTTACACCAACCTGGCC 435
/

```

```

512 GATCCTTTCTCTGAAGCTCTCTGGCTTCCCTGCTCTTGGACTTCTTCCACCATCTCCGG 571
QY 436 CAGAGATCTCTGAAGAGTTCACCTGCTGATGACACAGTGGGGAGGAGGCCAGGCTTC 495
DB 572 GAACAGATCCGAGAGAGTTCACCTGCTGATGACACAGTGGGGAGGAGGCCAGGCTTC 631
QY 496 CTGCGTGGTCTGCGGGTG-----AATGGGAGCCAGCCAGTACTTTTGGTGGTGTCCAT 549
DB 632 CTGGGTAGCTCCGCTGGCCGCGCAGAGGAGCCGCGCGGACCTTTGTCGGCGTCCAC 691
QY 550 GTGCGCGGAGGAGCTATGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 609
DB 692 GTGCGCGGTGGGAGCTATGTCAGAGTATGTCAGAGTATGTCAGAGTATGTCAGAGTAT 751
QY 610 CCGGGTTACTTGGAAGAGCCCTGGATATGTCGGGCGAGGCTATTCATCTCCAGTCTTC 659
DB 752 AGCGCTTACTCTCGGCGAGGCGCATGAGTGGTTCGGGCGAGGCGGACGAGCCCGCTTC 811
QY 670 GTGTTTACAAGCAACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
DB 812 GTGCTCACCAGCAACGCGCATGAGTGGTGTAAAGAAAACATCGACACCTCCAGGCGCAT 871
QY 730 GTGCTGCTCGCGGCAATGCTATGAGGGGTCCCGAGCCAGCAAGCACTTCGCGTCTCACC 789
DB 872 GTGACGTTTGTGGCGATGGACAGGAGGCTACACCGTGGAAAGACTTTGCCCTGCTCACA 931
QY 790 CAGTGCACACACACCATCATGACTATGAGACCTTTGGGATTTGGGCTGGCTACCTGGCA 849
DB 932 CAGTGCACACACACCATCATGACTATGAGACCTTTGGGATTTGGGCTGGCTACCTGGCT 991
QY 850 GGTGTGATACCATCTACTTAGCCAACTACACCCCTTCCGGATTCCTCGTTCCTCAAGTC 909
DB 992 GCGGAGAGACATGCTTACCTGCGCACTTCCAGCTGCGAGACTCTGAGTCTCTGAAGATC 1051
QY 910 TTTAGCCAGAGACGCTTCTTACCCGAATGGTGGGATCCCTCCGATCTGTCCTCCCA 969
DB 1052 TTTAGCCAGAGGCGCTTCTTCCCGGATGGTGGGATTAATGACAGACTTGTCTCCA 1111
QY 970 CTC 972
DB 1112 CTC 1114

```

RESULT 8

```

/ US-08-433-271-1
/ Sequence 1, Application US/08433271
/ Patent No. 5891698

```

GENERAL INFORMATION:

```

/ APPLICANT: Prieto, P., Smith, D., Cummings, R., Kopchick, J., Mukerji, P.,
/ APPLICANT: Moremen, K., Pierce, J.
/ TITLE OF INVENTION: Humanized Milk
/ NUMBER OF SEQUENCES: 1
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Donald O. Nicky
/ ADDRESSEE: ROSS Products Division
/ ADDRESSEE: Abbott Laboratories
/ STREET: 625 Cleveland Avenue
/ CITY: Columbus
/ STATE: Ohio
/ COUNTRY: United States of America
/ ZIP: 43215
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS Version 6.21
/ SOFTWARE: WordPerfect Version 6.0a
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/433,271
/ FILING DATE: 02-MAY-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/209,122
/ FILING DATE: 09-MAR-1994

```

CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614) 624-7080
TELEFAX: (614) 624-3074
TELEX: No. 5891698e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Cloned cDNA representing the product of a
MOLECULE TYPE: human genomic DNA segment
DESCRIPTION: GDP-L-fucose- D-galactoside 2'-alpha-fucosyl-transferase
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Entire amino acid sequence provided.
ORIGINAL SOURCE: Human Epidermal Carcinoma Cell line
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE: Human Epidermal Carcinoma Cell line
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 19
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: DNA sequencing and restriction analysis
OTHER INFORMATION: The encoded product of nucleotide SEQ ID NO:
OTHER INFORMATION: 1: is the enzyme, GDP-L-fucose- D-galactoside 2'-alpha-fucosyl-transferase
OTHER INFORMATION: which has the amino acid sequence described in SEQ. ID NO:1:
OTHER INFORMATION: is responsible for the synthesis of 2'-fucosyllactose.
PUBLICATION INFORMATION:
AUTHORS: V.P. Rajan, et al.
TITLE:
JOURNAL: Journal of Biological Chemistry
VOLUME: 264
ISSUE:
PAGES: 11158 - 11167
DATE: 1989
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-433-271-1

Query Match 38.2%; Score 408.2; DB 2; Length 1155;
Best Local Similarity 68.7%; Pred. No. 8.5e-111;
Matches 579; Conservative 0; Mismatches 258; Indels 6; Gaps 1;

QY 136 GGCATGTTCAAGTCAATTCATTTGGCCGGTGGGGAACAGATGGGGATACGCCACA 195
DB 272 GGCACCTGGAGCTGTCTACCCCAATGGCCGGTGGTAAATCAGATGGGACAGTGGCAG 331
QY 196 CTTCTTGATGGCCAGGATGACGACGCGTGGTTCATCCCGGCATCCATGGACACAC 255
DB 332 CTGCTGGCTCTGGCCAGCTCAACGCGCCGGGCTTTATCTGCTGCCATGATGCC 391
QY 256 GCTTAGGCCCATCTTTCAGGATCACTCCCGGTGTTACAGAGCAGACGCGCCAAAAG 315
DB 392 GCCCTGGCCCGGTATTCGGATCACTCCCGGTGTTACAGAGCAGACGCGCCAAAAG 451
QY 316 ATCCCATGGCAGATTACATCTCAACGACTGGATGGAGGACGTTACCGCCACATTCGG 375

DB 452 ACCCGTGGGGGAGCTGACGCTTACGACTGATGTCGAGAGGAGTACGGGACTGAGA 511
QY 376 GGACACTTTGTGGCTTCACGGGATACCCGTGCTCTGGACCTTCTACCAACCTGCGC 435
DB 512 GATCCTTTCTGTAAGCTCTCTGGCTTCCCTCTTGGACTTTCTCCACCATCTCCGG 571
QY 436 CCAGAGATCTGAAGGAGTTCACCTTGCATGACACGTCGGGAGGAGGCGCCAGCTTC 495
DB 572 GAACAGATCCGAGAGAGTTCACCTTGCACGACACCTTCGGAAGAGGCGCAGAGTGT 631
QY 496 CTGCTGGTCTCGGGGTG-----AATGGAGCAGCGGAGTACTTTGTGGTGTCCAT 549
DB 632 CTGGGTCACTCGCCCTGGCCGACAGGGGACGCGCCGACCTTTGTGCGGCTCCAC 691
QY 550 GTGCGCGGAGGAGTATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 609
DB 692 GTGCGCGGAGGAGTATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 751
QY 610 CGGGGTACCTGGAAAGGCGCTGATATGTCGCGGACGCTTATTCATCTCCAGTCTTC 669
DB 752 AGCGCTTACCTCGGAGGCGCTGATATGTCGCGGACGCTTATTCATCTCCAGTCTTC 811
QY 670 GTGTTTACAGCAACGCTATGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAC 729
DB 812 GTGTCACAGCAACGCTATGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAC 871
QY 730 GTGTTTACAGCAACGCTATGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAC 789
DB 872 GTGACGTTTGTGCGGAGAACATTAATGCTTCCGAGGAGAACATTAATGCTTCCGAG 931
QY 790 CAGTGCACACACACCATCATGATATGTCGCGGAGAACATTAATGCTTCCGAGGAG 849
DB 932 CAGTGCACACACACCATCATGATATGTCGCGGAGAACATTAATGCTTCCGAGGAG 991
QY 850 GTGTTTACAGCAACGCTATGCTGCTGCGGAGAACATTAATGCTTCCGAGGAGAC 909
DB 992 GCGGAGAGACATGCTTACCTGCGGAGAACATTAATGCTTCCGAGGAGAC 1051
QY 910 TTAAACCCAGAGCAGCCTTCCACCGGAGAACATTAATGCTTCCGAGGAGAC 969
DB 1052 TTAAACCCAGAGCAGCCTTCCACCGGAGAACATTAATGCTTCCGAGGAGAC 1111
QY 970 CTC 972
DB 1112 CTC 1114

RESULT 9

PCT-US91-00899-10

Sequence 10, Application PC/TUS9100899

GENERAL INFORMATION:

APPLICANT: Lowe, John B.

TITLE OF INVENTION: Method and Products For the Synthesis of

TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,

TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned

TITLE OF INVENTION: Genetic Sequences That Determine These Structures

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

ADDRESSEE: P.C.

STREET: 1755 Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00899

FILING DATE: 19910214

[illegible]

```
QY 256 GCCTAGCGCCCACTTCAGATCAGCCTCCCGTGTATACACAGCGACAGCGCCAAAAG 315
Db 393 GCCTAGCGCCCACTTCAGATCAGCCTCCCGTGTATACACAGCGACAGCGCCAAAAG 452
QY 316 ATCCATGCGCAGATACACATCAGCCTCCCGTGTATACACAGCGACAGCGCCAAAAG 375
Db 453 ACCTGCGCGCGGAGTGCAGCTTCAGCCTCCCGTGTATACACAGCGACAGCGCCAAAAG 512
QY 376 GGACACTTTTGCCTTCAGGATACCGGATACCGGATACCGGATACCGGATACCGGATAC 435
Db 513 GATCCTTTCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
QY 436 CCAGAGATCTGAGGAGTTCACCTGCTATACACAGCGACAGCGCGGAGGAGCGCGGCTTC 495
Db 573 GAACAGATCCGACAGAGTTCACCTGCTATACACAGCGACAGCGCGGAGGAGCGCGGCTTC 632
QY 496 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
Db 633 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
QY 609 GTGCGCGCGGAGGACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
Db 753 AGCGCTTACCTGCGGAGGAGGACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 812
QY 729 GTGCTTACAGCAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
Db 813 GTGCTTACAGCAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872
QY 789 GTGCTTACAGCAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
Db 933 CAGTGAACACACACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 992
QY 909 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052
Db 993 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052
QY 910 TTAAAGCGAGGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
Db 1053 TTAAAGCGAGGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1112
QY 970 CTC 972
Db 1113 CTC 1115
```

RESULT 11

PCT-US95-07554-3

Sequence 3, Application PC/US9507554

GENERAL INFORMATION:

APPLICANT: Sandrin, Mauro S.

APPLICANT: Fodor, William L.

APPLICANT: Rother, Russell P.

APPLICANT: Squinto, Stephen P.

APPLICANT: McKenzie, Ian F. C.

TITLE OF INVENTION: Methods for Reducing

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

STREET: 1951 Burr Street

CITY: Fairfield

STATE: Connecticut

COUNTRY: USA

ZIP: 06430

COMPUTER READABLE FORM:

```
MEDIUM TYPE: 3.5 inch, 750 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07554
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,201
FILING DATE: June 15, 1994
CLASSIFICATION:
APPLICATION NUMBER: 08/278,282
FILING DATE: July 21, 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-144.1PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TCPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Human H-transferase
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Larsen, R.D.
AUTHORS: Ernst, L.K.
AUTHORS: Nair, R.P.
AUTHORS: Lowe, J.B.
TITLE: Molecular cloning, sequence, and
expression of a human GDP-L-fucose:
TITLE: -D-galactoside 2-alpha-L-
TITLE: fucosyltransferase cDNA that can
TITLE: form the H blood group antigen.
JOURNAL: Proceedings of the National
Academy of Sciences, USA
VOLUME: 87
PAGES: 6674-6678
DATE: SEP-1990
PCT-US95-07554-3
```

Query Match 37.9%; Score 405; DB 5; Length 1174;

Best Local Similarity 68.4%; Pred. No. 7.5e-110;

Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

```
QY 136 GGCATGTTACGATCAATTCATTTGGCGGCTGGGACACAGATGGCGCAATACGCCACA 195
Db 249 GGCACCTGGAGTCTTACCCCAATGGCGGCTTTGGTAAATCAGATGGACAGTAIGCCACG 308
QY 196 CTCTTTGCATGCGCAGGATGACGACGCGTTCATCCCGCATCCATGCGACAC 255
Db 309 CTGCTGGCTCTGGCCAGCTCAACGCGCGGCGCTTTATCTGCTGCTGCTGCTGCTGCTG 368
QY 256 GCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTGTACACAGCGACAGCGCCAAAAG 315
Db 369 GGCCTGGCCCCGTATTCGCGATCACCCCTGCGCTGCTGCGCCCAAGATGGACACCGC 428
QY 316 ATCCCATGGCAGATATACCATCTCAACGAGTGGAGGAGCGTTACCGCCACATTCGG 375
Db 429 ACGCCGTGGCGGAGCTGCGAGCTTCACGACTTCGCGAGGAGTACGCGGAGCTTGAGA 488
QY 376 GGACACTTTTGGCGCTTCACGGGATACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
Db 489 GATCCTTTCTTGAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548
```



```

QY 970 CTC 972
Db 1125 CTC 1127

RESULT 13
US-08-675-773B-4
; Sequence 4, Application US/08675773B
; Patent No. 6166288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; TITLE OF INVENTION: FOR XENOTRANSPLANTATION.
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESS: BROWDY AND NEWMARK, P.L.L.C.
; STREET: 419 SEVENTH STREET, N.W., SUITE 300
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/675,773B
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: DIAMOND-1A
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-675-773B-4

Query Match 37.9%; Score 405; DB 4; Length 2268;
Best Local Similarity 68.4%; Pred. No. 1e-105;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCAGTTTACAGATATCCATTGCGCGCTGGGGAACAGATGGGGAATACGCCACA 195
Db 749 GGCACCTGGAGCTCTACACCAATGCGCGTTTGGTAATCAGATGGGACAGTATGCCAG 808

QY 196 CTTCTTGGCACTGGCCAGGATGACGACGCGCTTGCCTTATCCCGCATCCATGACAC 255
Db 809 CTGCTGGCTTGGCCAGCTACACGCGCCGCGGCTTTATCTGCTGCCATGCTGCG 868

QY 256 GTCTAGCGCCCATCTTACAGATACAGCTTCCAGTGGAGGAGGCTTACCGCCCATTCGG 315
Db 869 GCGCTTGGCCCGGATATCCGCTATCACCCTTGGCGCTGCTGCGCCAGAGTGGACAGCCGC 928

QY 316 ATCCCATGGCAGAAATACCATCTCAACGACTGGATGGAGGAGGCTTACCGCCCATTCGG 375
Db 929 ACGCCGTGGCGGAGCTGACGCTTACAGACTGGATGCTGGAGGAGTACCGGACITGAGA 988

QY 376 GACACATTTTGGGCTTACCGGATACCGGCTCTCTGACCTTTCACCAACACCTTGGC 435
Db 989 GATCCTTCTGAAAGCTCTGCTGCTTCCCTTCCCTTGGACTTCTTCCACCATCTCCGG 1048

```

```
TELEPHONE: (404) 815-6508
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3373 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 104..1201
OTHER INFORMATION: /note= "Nucleotides 104 through 1201 encode the GDP-L-fuc
PUBLICATION INFORMATION:
AUTHORS: Larsen, et al.
JOURNAL: Proc. Nat'l Acad. Sci. USA
VOLUME: 87
PAGES: 6674-6678
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
US-08-273-411-2

Query Match 37.9%; Score 405; DB 1; Length 3373;
Best Local Similarity 68.4%; Pred. No. 1.2e-109;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCATGTTTCACGATCAATTCATTCGCGGCTGGGGAACACAGATGGGGAATACGCCACA 195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 196 GTCTTTGACATGCGCCAGGATGAACGGACGGCTTGCCTTCATCCCGCATCCATGCACAAAC 255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 398 CTGCTGGCTCTGCGCCAGCTCAACGGCGCGCGGCTTATCTGCTGCCATGATGCC 457
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 GCTCTAGCGCCCATCTTACGATCAGCTCCCGGTGTTACAGGACGACAGGCCCAAAAG 315
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 458 GCGCTGGCGCGGATTCGCGCATACCTGCGCGTCTGGCGCCAGAGTGGACAGCGCC 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 316 ATCCCATGTGCAGAAATACCATCTCAACGACTGGATGGAGGCGTTACGCCACATCGG 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 518 ACGCGTGGCGGAGCTGCAGCTTACGACTGATGTCGAGAGAGTACCGGACTTGAGA 577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 GGACACTTTTGGGCTTCAGGGATACCGCTGCTCTGGACCTTACACCAACCTGCGC 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 578 GATCCTTTTCTGAAGCTCTGCTGCTTCCCTGCTCTGGACTTCTTCCACCATCTCCGG 637
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 436 CCAGAGATCCTGAAGAGTTACCTGATGACCATGACCGTCCGGAGGAGGCCAGGCCCTTC 495
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 638 GAACAGATCCGAGAGTTACCTGACACGACCATCTTCGGGAGAGGCGGAGAGTGTG 697
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 496 CTGCGTGTCTCGGGTG-----AATGGAGCCAGCGGACTATTTTGTGGGTGTCCAT 549
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 698 CTGGGTACGCTCCGCTGGCGGCACAGGGAGCCGCCGCCACCTTTCTGCGCGTCCAC 757
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 550 GTGCGCCGAGGGACTATGTGATGTCATGCTATGCTGGAAGGCGCTGGTGGCGTAC 609
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 758 GTGCGCGGTGGGACTATCTGAGGTTATGCTGAGCTGCTGAGGAGGTTGTTGGGCGAC 817
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 610 CGGGGTTACCTGGAAAGGCCCTGGATATGTTCCGGGCGAGCTATTTCATCTCCAGTCTTC 669
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 818 AGCGCTACTCTCGGCGAGCGATGGAGTGTGGTCCGGGCGAGCGACGAAACCCCGCTTTC 877
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 GTGGTTACAGCAACGGTATGGCTGTGGTCCGGGAGAGCAATTAACTCTCCCGAGGAGAC 729
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 878 GTGTCACCAAGCAACGGCATGGAGTGTGTAAGAAACATCCACACTTCCAGGCGCAT 937
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 730 GTGGTTCGCGGGCAATGTATGTAGGGGTGCCAGCCAAAGGACTTCGGCTGTCTACC 789
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 938 GTACGCTTTGCTGGCGATGACAGGAGGCTACACCGTGGAAAGACTTTGCCCTGTCTACA 997
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 790 CAGTGCACACACCATCATGACTATTGGGACCTTTGGGATTTGGGCTGCTACCTACCTGGCA 849
```

```
Db 998 CAGTGCACACACCATCATGACTATTGACCATTTGGACCTTCTGGGCTTCTGGGCTGCCTACCTGGCT 1057
QY 850 GTGTGTATACCATCTACTTACCGAATACACCTTCCGGATTCCTCCGTTCTCTCAAGTC 909
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1058 GCGGAGAGACTGCTACTTGGCCAACTTCACTTCCGAGACTGAGTTCCTGAGATC 1117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 TTAAAGCCAGAGGAGCCCTTCTACCCGAATGGTGGGATCCCTCGGATCTTCCCA 969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1118 TTAAAGCCGAGGCGCTTCTCTGCCGAGTGGTGGGCAATTAATGCAGACTTGTCTCCA 1177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 970 CTC 972
Db |||||
QY 1178 CTC 1180
Db |||||

RESULT 15
US-08-675-773B-3
; Sequence 3, Application US/08675773B
; Patent No. 6166288
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, LISA E
; APPLICANT: LOGAN, JOHN S
; APPLICANT: BYRNE, GUERARD W
; APPLICANT: SHARMA, AJAY
; TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 SEVENTH STREET, N.W., SUITE 300
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,773B
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: DIAMOND=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3791 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-675-773B-3

Query Match 37.9%; Score 405; DB 4; Length 3791;
Best Local Similarity 68.4%; Pred. No. 1.3e-109;
Matches 577; Conservative 0; Mismatches 260; Indels 6; Gaps 1;

QY 136 GGCATCTTCACGATCAATTCATTCGCGGCTGGGGAACACAGATGGGGAATACGCCACA 195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2293 GGCACCTGCACTCTCTACCCCAATGCCGCTTGGTAATCAGATGGAGACAGTATGCCAGC 2352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 196 CTCTTTTGCATGCGCCAGGATGAACGGACGGCTTGCCTTCATCCCGCATCCATGCACAAAC 255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2353 CTGCTGGCTCTGGCCAGCTCAACGGCGCGGCTTATCTCTGCTGCCATGCTGCC 2412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 GCTCTAGCGCCCATCTTCAGGATCAGCTCCCGGTTTACACAGCAGCGGCCCAAAAG 315
```

Db 2413 GCCCTGGCCCCGGTATTCCGCATCACCCCTGCCGTCGCCCGACGAAGTGGACACCCGC 2472
QY 316 ATCCCATGCAAGAAATTACCATCTCAACGACTGGATGGAGAGCGTTACCGCCACATTCGG 375
Db 2473 AGCCCTGGCGGAGCTGAGCTTCACGACTGGATGTCGAGGAGTACCGGACTTGAGA 2532
QY 376 GGACACTTTGTGGCTTCACGGGATACCCGTGCTGCTGACCTTCTACCAACACCTGGCG 435
Db 2533 GATCCTTTCTGAAGCTCTGCTGGCTCCCTGCTCTTGGACTTTCTCCACCACTCCCG 2592
QY 436 CCAGAGATCCTGAAGGAGTCAACCTGTCATGACACGTCGCGGAGGAGGCCAGGCTTC 495
Db 2593 GAACAGATCGCAGAGAGTTCACCTGTCACGACCACTTCGCGGAGAGCGCAGAGTG 2652
QY 496 CTGCTGGTCTGGGGTG-----AATGGAGCAGCCGAGTACTTTTGGGTGTCCAT 549
Db 2653 CTGGGTCAAGCTCGGCTGGCGCGCACAGGGGACCGCCGCGCACCTTTGTGGCGTCCAC 2712
QY 550 GTGCGCGGAGGAGTATGTGCATGTCTGCTAAATGTGTGGAGGCGTGGTGGCTGAC 609
Db 2713 GTGCGCGTGGGACTATCTGCAGGTTATGCCCTCAGCGCTGGAGGCTGGTGGCGGAC 2772
QY 610 CGGGTTACTGGAAAGGCCCTGGATATGTCGGGCAAGCTTATTCATCTCCAGTCTTC 669
Db 2773 AGCGCTTACTCGGCGAGCCATGGACTGGTTCGGGCGACGGCAAGGCCCGCTTTC 2832
QY 670 GTGGTTACAGCAACGCTATGSCCTGGTGGCGGAGAACATTAATGCTTCCGAGGAGAC 729
Db 2833 GTGGTCACGAGCAACGGCATGGAGTGGTAAAGAAACATCGACACCTCCCAAGGCGAT 2892
QY 730 GTGGTGTTCGGGGCAATGGTATTGAGGGGTGCCAGCAAGGACTTCGGCTGCTCAAC 789
Db 2893 GTGACGTTTGTGGCGATGGACAGGAGGCTACACCGTGAAGACTTTGCCCTGCTCACA 2952
QY 790 CAGTGAACACACCATCATGACTATTGGAGCTTTGGGATTTGGCTGCTACCTGGCA 849
Db 2953 CAGTGAACACACCATTTATGACATTTGGACCTTCGGCTTCTGGGCTGCCCTACCTGGCT 3012
QY 850 GGTGGTGATACCATCTACTAGCCACTACACCCCTCCGGATTTCTCCGTTCTCAAGTC 909
Db 3013 GCGGAGACACTGTCTACCTGGCCAACTTCACCCCTGCCAGACTCTGAGTTCTCTGAAGTC 3072
QY 910 TTAAAGCCAGAGCAGCTTCCTACCCGAATGGGTGGGATCCCTGCCGATCTGTCCCA 969
Db 3073 TTAAAGCCGAGGCGGCGCTTCTGCGCGAGTGGGTGGGATTAATGACAGACTGTCTCCA 3132
QY 970 CTC 972
Db 3133 CTC 3135

Search completed: May 26, 2003, 11:58:34
Job time : 60.2896 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 11:58:37 ; Search time 163.789 Seconds
(without alignments)
8610.191 Million cell updates/sec

Title: US-10-040-863-9

Perfect score: 1068

Sequence: 1 CTCACGAGCAATAGTGGAA.....GAAGAGCTCTGTGATGGAA 1068

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:

```
1: /cgn2_5/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_5/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_5/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_5/ptodata/1/pubpna/PC1US_PUBCOMB.seq:*
7: /cgn2_5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_5/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_5/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_5/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_5/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_5/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1068	100.0	1068	10	US-09-999-672-9
2	1068	100.0	1068	12	US-10-040-863-9
3	1068	100.0	1068	10	US-09-999-672-7
4	1068	100.0	1068	12	US-10-040-863-7
5	612.8	57.4	1043	10	US-09-051-034A-1
6	589.4	55.2	3088	10	US-09-954-456-45
7	589.4	55.2	3088	10	US-09-954-456-121
8	589.4	55.2	3088	10	US-09-969-347-234
9	423.8	39.7	1098	10	US-09-051-034A-3
10	423.8	39.7	1269	10	US-09-844-268-12
11	423.8	39.7	1269	10	US-09-844-705-12
12	405	37.9	3373	9	US-10-105-963-9
13	405	37.9	8174	10	US-09-863-475A-5
14	72.8	6.8	100	10	US-09-999-672-29
15	72.8	6.8	100	12	US-10-040-863-29
16	71.2	6.7	100	10	US-09-999-672-23
17	71.2	6.7	100	10	US-09-999-672-26
18	71.2	6.7	100	12	US-10-040-863-23
19	71.2	6.7	100	12	US-10-040-863-26

20	69.6	6.5	100	10	US-09-999-673-20	Sequence 20, Appl	
21	69.6	6.5	100	12	US-10-040-863-20	Sequence 20, Appl	
22	63.2	5.9	100	10	US-09-999-673-14	Sequence 14, Appl	
23	63.2	5.9	100	10	US-09-999-672-17	Sequence 17, Appl	
24	63.2	5.9	100	12	US-10-040-863-14	Sequence 14, Appl	
25	63.2	5.9	100	12	US-10-040-863-17	Sequence 17, Appl	
c	26	40.4	3.8	3044	9	US-10-125-540-98	Sequence 98, Appl
c	27	40.4	3.8	3044	10	US-09-764-870-98	Sequence 98, Appl
28	38.4	3.6	50	10	US-09-999-672-16	Sequence 16, Appl	
29	38.4	3.6	50	12	US-10-040-863-16	Sequence 16, Appl	
30	37.6	3.5	1569	9	US-10-051-643-113	Sequence 113, App	
31	37.6	3.5	1569	9	US-09-880-505-113	Sequence 113, App	
32	37.6	3.5	1626	9	US-10-051-643-159	Sequence 159, App	
33	37.6	3.5	1626	9	US-09-880-505-159	Sequence 159, App	
34	37.4	3.5	50	10	US-09-999-672-28	Sequence 28, Appl	
35	37.4	3.5	50	12	US-10-040-863-28	Sequence 28, Appl	
36	37	3.5	50	10	US-09-999-672-25	Sequence 25, Appl	
37	37	3.5	50	12	US-10-040-863-25	Sequence 25, Appl	
38	36.8	3.4	50	10	US-09-999-672-19	Sequence 19, Appl	
39	36.8	3.4	50	12	US-10-040-863-19	Sequence 19, Appl	
40	36.4	3.4	802	9	US-10-184-634-312	Sequence 312, App	
41	36.4	3.4	802	9	US-10-184-634-312	Sequence 312, App	
c	42	36.4	3.4	2030	10	US-09-880-107-2416	Sequence 2416, App
43	36.2	3.4	916	9	US-10-184-644-78	Sequence 78, Appl	
44	36.2	3.4	916	9	US-10-184-644-78	Sequence 78, Appl	
45	36	3.4	1416	8	US-08-900-220C-19	Sequence 19, Appl	

ALIGNMENTS

RESULT 1

US-09-999-672-9

; Sequence 9, Application US/09999672

; Patent No. US20020127655A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHAL-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/399,672

; CURRENT FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: US/09/298,886

; PRIOR FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 1068

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1062)

; US-09-999-672-9

Query Match 100.0%; Score 1068; DB 10; Length 1068;

Best Local Similarity 100.0%; Pred. No. 6;

Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGAGCAATAGTGGAAACACAGAAAGCCAGAGATCGGCGGACAGCAGCATGGG 120

Db 1 CTCACGAGCAATAGTGGAGCTCCACCCCTCTCAGAGAGAAATACCGATGAGCACT 60

QY 61 CAAATGCTCTCGGAAACACAGAAAGCCAGAGATCGGCGGACAGCAGCATGGG 120

Db 61 CAAATGCTCTCGGAAACACAGAAAGCCAGAGATCGGCGGACAGCAGCATGGG 120

QY 121 AATGGAGACTCGGGGGCATGTTCACAGATCAATTCATTGGCGGGAACACAGATG 180

Db 121 AATGGAGACTCGGGGGCATGTTCACAGATCAATTCATTGGCGGGAACACAGATG 180

QY 181 GGCGAATACGCCACACTCTTTCAGCTGGCCAGGATGACGAGCGGCTTGGCTTCATCCCC 240

181 GCGCAATACCCACACATCTTTGGCACTGGCCAGATGACCGAGCGCTTGGCTTCATCCCC 240
241 GCATCATGACAAACGCTCTAGCGCCCATCTTCCAGATCAGCCTCCCGGTGTTACACAGC 300
241 GCATCATGACAAACGCTCTAGCGCCCATCTTCCAGATCAGCCTCCCGGTGTTACACAGC 300
301 GACAGGCCAAAAGATCCCATGGCAGATTAACCATCTCAACACTGGATGGAGGAGCGT 360
301 GACAGGCCAAAAGATCCCATGGCAGATTAACCATCTCAACACTGGATGGAGGAGCGT 360
361 TACCGCCACATTCGGGACACTTTGTGCGCTTCCACGGATACCGGTCTCCTCGACCTTC 420
361 TACCGCCACATTCGGGACACTTTGTGCGCTTCCACGGATACCGGTCTCCTCGACCTTC 420
421 TACACACCTGCGGCCAGAGATCTCTGAAGAGTTCACCTGTCATGACCAACGCTGCGGAG 480
421 TACACACCTGCGGCCAGAGATCTCTGAAGAGTTCACCTGTCATGACCAACGCTGCGGAG 480
481 GAGGCCAGGCTTCTGCGTGGTCTGCGGTGAATGGGAGCCAGCGAGTCTTTTGTG 540
481 GAGGCCAGGCTTCTGCGTGGTCTGCGGTGAATGGGAGCCAGCGAGTCTTTTGTG 540
541 GGTGTCATCTGCGCGAGGAGTATGTGCAATGTCATGCTGCTGCTGCTGCTGCTGCTGCTG 600
541 GGTGTCATCTGCGCGAGGAGTATGTGCAATGTCATGCTGCTGCTGCTGCTGCTGCTGCTG 600
601 GTGGCTGACGGGTTACCTGGAAGGCGCTGGATGATGTTGCGGCAGCCTATTCATCT 660
601 GTGGCTGACGGGTTACCTGGAAGGCGCTGGATGATGTTGCGGCAGCCTATTCATCT 660
661 CCAGTCTTCTGTTTACAAGCAACGCTATGGCTGCTGCGGAGAACATTAATGCTTC 720
661 CCAGTCTTCTGTTTACAAGCAACGCTATGGCTGCTGCGGAGAACATTAATGCTTC 720
721 CGAGGAGAGTGGTGTTCGCGGCAATGATGAGGGTTCGCGAGCCAGGACTTCGCG 780
721 CGAGGAGAGTGGTGTTCGCGGCAATGATGAGGGTTCGCGAGCCAGGACTTCGCG 780
781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCCAGGACTTCGCG 840
781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCCAGGACTTCGCG 840
841 TACCTGGCAGTGGTGTATACCATCTACTTAGCAACTACACCTTCGCGATTCGCTTC 900
841 TACCTGGCAGTGGTGTATACCATCTACTTAGCAACTACACCTTCGCGATTCGCTTC 900
901 CTCAGAGTCTTTAAGCCAGAGGAGCCTTCTACCGAATGGGTGGGATTCGCGAT 960
901 CTCAGAGTCTTTAAGCCAGAGGAGCCTTCTACCGAATGGGTGGGATTCGCGAT 960
961 CTGTCCTCCACTCTTAAGCATTAAACACAGCCTGCTGCGGCTCCACTCCACCTCAAG 1020
961 CTGTCCTCCACTCTTAAGCATTAAACACAGCCTGCTGCGGCTCCACTCCACCTCAAG 1020
1021 GCAGAGGAGTCACTTGTACGTGCGAGGAGCCTTCTGATGGAA 1068
1021 GCAGAGGAGTCACTTGTACGTGCGAGGAGCCTTCTGATGGAA 1068

RESULT 2

US-10-040-863-9
; Sequence 9, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,986

; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
US-10-040-863-9

Query Match 100.0%; Score 1068; DB 12; Length 1068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCACGAGGATAGTGAAGCTCCACCCCTGTCAGAGAAGGAATACCGATACGACT 60
DB 1 CTCACGAGGATAGTGAAGCTCCACCCCTGTCAGAGAAGGAATACCGATACGACT 60
QY 61 CAATATGCTCTCGGAAACACAGAAAGCCAGAGATGCGAGCGGACACGAGCAGCAT 120
DB 61 CAATATGCTCTCGGAAACACAGAAAGCCAGAGATGCGAGCGGACACGAGCAGCAT 120
QY 121 AATGGAGAGCTGCGGGCATGTTACAGATCAATTCATTTGCGCGGCTGGGGAACAGATG 180
DB 121 AATGGAGAGCTGCGGGCATGTTACAGATCAATTCATTTGCGCGGCTGGGGAACAGATG 180
QY 181 GCGCAATACCCACACTCTTTGCACTGGCGAGATGAACGAGCGGTGCTTCATCCC 240
DB 181 GCGCAATACCCACACTCTTTGCACTGGCGAGATGAACGAGCGGTGCTTCATCCC 240
QY 241 GCATCATGACAAACGCTCTAGCGCCCATCTTCCAGATGATGAGCGGTGTTACACAGC 300
DB 241 GCATCATGACAAACGCTCTAGCGCCCATCTTCCAGATGATGAGCGGTGTTACACAGC 300
QY 301 GACAGGCCAAAAGATCCCATGGCAGATTAACCATCTCAACACTGGATGGAGGAGCGT 360
DB 301 GACAGGCCAAAAGATCCCATGGCAGATTAACCATCTCAACACTGGATGGAGGAGCGT 360
QY 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCCACGGATACCGGTCTCCTCGACCTTC 420
DB 361 TACCGCCACATTCGGGACACTTTGTGCGCTTCCACGGATACCGGTCTCCTCGACCTTC 420
QY 421 TACCGCCACCTGCGGCCAGAGATCCTGAAGAGTTCACCTGTCATGACCAACGCTGCGGAG 480
DB 421 TACCGCCACCTGCGGCCAGAGATCCTGAAGAGTTCACCTGTCATGACCAACGCTGCGGAG 480
QY 481 GAGGCCAGGCTTCTGCGTGGTCTGCGGTGAATGGGAGCCAGCGAGTACTTTTGTG 540
DB 481 GAGGCCAGGCTTCTGCGTGGTCTGCGGTGAATGGGAGCCAGCGAGTACTTTTGTG 540
QY 541 GGTGTCATCTGCGCGAGGAGGACTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 GGTGTCATCTGCGCGAGGAGGACTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 GTGGCTGACCGGTTACCTGGAAGGCGCTGGATGATGTTGCGGCAGCCTATTCATCT 660
DB 601 GTGGCTGACCGGTTACCTGGAAGGCGCTGGATGATGTTGCGGCAGCCTATTCATCT 660
QY 661 CCAGTCTTCTGTTTACAAGCAACGCTATGGCTGCTGCGGAGAACATTAATGCTTC 720
DB 661 CCAGTCTTCTGTTTACAAGCAACGCTATGGCTGCTGCGGAGAACATTAATGCTTC 720
QY 721 CGAGGAGAGTGGTGTTCGCGGCAATGATGAGGGTTCGCGAGCCAGGACTTCGCG 780
DB 721 CGAGGAGAGTGGTGTTCGCGGCAATGATGAGGGTTCGCGAGCCAGGACTTCGCG 780
QY 781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCCAGGACTTCGCG 840
DB 781 CTGCTCAGCCAGTGCACACACACATCATGATGATGAGGGTTCGCGAGCCAGGACTTCGCG 840
QY 841 TACCTGGCAGTGGTGTATACCATCTACTTAGCAACTACACCTTCGCGATTCGCTTC 900

Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3
AAB36104
ID AAB36104 standard; Protein; 380 AA.
XX
AC AAB36104;
XX
DT 19-FEB-2001 (first entry)
XX
DE Rat hepatoma H35 cell alpha1-2fucosyltransferase.
XX
KW Rat; alpha1-2fucosyltransferase; cytostatic; neuroprotective;
KW neotropic; gene therapy; Fucalbeta1-2Galbeta1-3GalNAc; immunotherapy;
KW immunosuppression; cancer; neurological disease;
KW small cell lung carcinoma.
XX
OS Rattus norvegicus.
XX
PN WO200064464-A1.
XX
PD 02-NOV-2000.
XX
PF 23-APR-1999; 99WO-US07384.
XX
PR 23-APR-1999; 99WO-US07384.
XX
PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.
XX
PI Holmes EH, Sherwood AL;
XX
DR WPI; 2000-687262/67.
DR N-PSDB; AAG67965.
XX
XX
PT New rat ganglioside GM1-specific alpha1-2Fucosyltransferase, useful for
PT preparation of fucosyl GM1 which is useful as a nutritional composition
PT or immunotherapeutic for cancer and neurological diseases -
XX
PS Claim 1; Fig 5; 91pp; English.
XX
XX The present sequence was given in a specification relating to an isolated
CC rat ganglioside GM1-specific alpha1-2fucosyltransferase protein. The
CC protein or its cellular fraction is useful for synthesis of a molecule
CC comprising Fucalbeta1-2Galbeta1-3GalNAc, a glycolipid, glycoprotein,
CC glycolipoprotein or a free oligosaccharide comprising
CC Fucalbeta1-2Galbeta1-3GalNAc. The method involves contacting
CC alpha1-2fucosyltransferase with GDP-fucose and a molecule or glycolipid,
CC galactose, glycolipoprotein or oligosaccharide having a terminal
CC Galbeta1-3GalNAc group. It is also useful for synthesis of fucosyl-GM1
CC by contacting the protein with GDP-fucose and ganglioside GM1. The
CC obtained glycoproteins, glycolipoproteins, glycolipids and
CC oligosaccharides are useful as nutritional compositions and fucosyl-GM1
CC is useful for inducing an immunotherapeutic or immunosuppressive action
CC against cancer, neurological disease or small cell lung carcinoma.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1896; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-198;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKLQPLSEKELPMTTQMSSGNTSEPMRDRDSQHGNGELGRGFTINSIGRLGNQM 60
Db 28 LQQRIVKLQPLSEKELPMTTQMSSGNTSEPMRDRDSQHGNGELGRGFTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASHNNAAPFRISLPVLSHSDTAKKIPQWNYHLNDWMEER 120
Db 88 GEYATLALARMNGRLAFIPASHNNAAPFRISLPVLSHSDTAKKIPQWNYHLNDWMEER 147
QY 121 YRHIPGHVFTGYPCSWTTHLRLPEILKEFTLHDHVREAAFLRLGVNGSQPSITFV 180
Db 148 YRHIPGHVFTGYPCSWTTHLRLPEILKEFTLHDHVREAAFLRLGVNGSQPSITFV 207

QY 181 GVHVRGDIYVHMPNVWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 208 GVHVRGDIYVHMPNVWKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 267
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTTIGTFGFWAAYLAGGDTIYLYNTLPDSPF 300
Db 268 RGDVVFAGNGIEGSPAKDFALLTQCNTHTTIGTFGFWAAYLAGGDTIYLYNTLPDSPF 327
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353
Db 328 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380

RESULT 4
AAE16622
ID AAE16622 standard; Protein; 380 AA.
XX
AC AAE16622;
XX
DT 09-APR-2002 (first entry)
XX
DE Rat hepatoma H35 cell alpha1-2FUCT.
XX
KW Rat; alpha1-2fucosyltransferase; alpha1-2FUCT; antisense therapy;
KW galactose beta1-3N-acetylgalactosamine; Galbeta1-3GalNAc; glycolipid;
KW glycoprotein; glycolipoprotein; oligosaccharide; fucosyl-GM1; cancer;
KW gene therapy; oncogenic transformation; cytostatic; ganglioside; GM1;
KW cell transformation.
XX
OS Rattus norvegicus.
XX
XX
FH Key Location/Qualifiers
FT Domain 1..27 /note= "Intracellular/Transmembrane domain"
FT Modified-site 199 /note= "N-glycosylated"
FT Region 221..380
FT Modified-site 265 /note= "Region which overlaps rat FTEB"
FT Modified-site 293 /note= "N-glycosylated"
FT Modified-site 293 /note= "N-glycosylated"
XX
XX US6329170-B1.
XX 11-DEC-2001.
XX 23-APR-1999; 99US-0298886.
XX 23-APR-1999; 99US-0298886.
XX (NWHO-) NORTHWEST HOSPITAL.
XX Holmes EH, Sherwood AL;
XX
XX WPI; 2002-121132/16.
XX N-PSDB; AAD27207.
XX
XX Rat hepatoma H35 cell alpha1-2fucosyltransferase, useful for producing
XX GM1-specific alpha1-2fucosyltransferase enzyme by recombinant
XX techniques and for detecting oncogenic transformation of test tissues -
XX
XX Claim 3; Fig 5; 41pp; English.
XX
XX The invention relates to rat GM1-specific alpha1-2fucosyltransferase
XX (alpha1-2FUCT) enzyme and its corresponding nucleic acid. This nucleic
XX acid is specific for a carbohydrate moiety found in ganglioside GM1,
XX a terminal galactose beta1-3N-acetylgalactosamine (Galbeta1-3GalNAc),
XX saccharide. Alpha1-2FUCT DNA is useful for producing rat alpha1-2FUCT
XX protein by recombinant techniques. Alpha1-2FUCT DNA is useful for the
XX preparative synthesis of fucosyl containing glycolipids, glycoproteins,
XX glycolipoproteins and oligosaccharide, and for preparing fucosyl-GM1.

DT 08-JUL-1998 (first entry)
XX Pig secretor.
XX Pig; secretor; chimeric; glycosyltransferase; gene therapy;
KW transplantation.
XX Sus scrofa.
XX WO9805768-A1.
FN 12-FEB-1998.
PD
XX 01-AUG-1997; 97WO-AU00492.
XX 21-AUG-1996; 96US-0024279.
PR 02-AUG-1996; 96AU-0001402.
XX (AUST-) AUSTIN RES INST.
PA McKenzie IFC, Sandrin MS;
XX WPI; 1998-159170/14.
DR N-PSDB; AAV21639.
XX Nucleic acids encoding chimeric glycosyltransferases - used for
PT altering carbohydrate levels on the surface of cells, useful in gene
PT therapy and transplantation
XX Example 2; Fig 6; 51pp; English.
XX The present sequence represents pig secretor used in an example of the
CC present invention. The present invention describes nucleic acids (NA)
CC encoding a chimeric glycosyltransferase. The NA comprises a catalytic
CC domain of a first glycosyltransferase (GT) and a localisation signal of
CC a second GT, whereby when the NA is expressed in a cell and where the
CC chimeric enzyme is located in an area of the cell where it is able to
CC compete for substrate with a second GT, resulting in reduced levels of
CC a product from the second GT. The NA can be used to produce cells and
CC organs with desired glycosylation patterns. Products and methods of the
CC present invention can be used to reduce the levels of undesirable
CC epitopes in cells, tissues or organs which may be used in
CC transplantation or gene therapy.
XX Sequence 340 AA;
Query Match 75.4%; Score 1430; DB 19; Length 340;
Best Local Similarity 81.9%; Pred. No. 1.2e-147;
Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;
QY 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPMPRRDSEQHNGELRGMTINSIGRLGNQM 60
Db LQORIVKLQPLSEKELPMTTOMSSGNTESPMPRRDSEQHNGELRGMTINSIGRLGNQM 73
QY 61 GEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWNYHLNDWMEER 120
Db 74 GEYATLALARMNGRPAFIPPEMSTLAFIPRITLPVLHASTARRIPWNYHLNDWMEER 133
QY 121 YRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
Db 134 YRHIPGEYVRLTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSRPSTFV 193
QY 161 GVHVRGDYVHVMNVKGVYVADRGYLEKALDMFRARYSPFVVTSGNAWCNENAS 240
Db 194 GVHVRGDYVHVMNVKGVYVADRGYLEKALDMFRARYSPFVVTSGNAWCNENAS 253
QY 241 RGVVVFAGNTEGSPAKDFALLTQCNHTIMTIGTFGWAAYLAGDFTIYLYNTLPDPSF 300
Db 254 RGVVVFAGNTEGSPAKDFALLTQCNHTIMTIGTFGWAAYLAGDFTIYLYNTLPDPSF 313
QY 301 LKVEKPEAAFLPEWVGIPADLSPLK 326
Db 314 LKLFKPEAAFLPEWVGIEADLSPLK 339

RESULT 7
AAW69332
ID AAW69332 standard; P:Protein; 343 AA.
XX AC AAW69332;
DT 20-NOV-1998 (first entry)
DE Human Sec2 protein sequence.
XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
KW FUT2; nonsecretor genotyping.
OS Homo sapiens.
XX US5807732-A.
PD 15-SEP-1998.
XX 28-FEB-1995; 95US-0395800.
XX 28-FEB-1995; 95US-0395800.
XX (GIOR/) GIORGI D.
PA (KELL/) KELLY R. J.
PA (LENN/) LENNON G.
PA (LOWE/) LOWE J. B.
PA (ROUQ/) ROUQUIER S.
XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
WPI; 1998-520127/44.
DR N-PSDB; AAV58323.
XX DNA encoding fucosyltransferase enzyme - useful for producing
PT recombinant enzyme and genotyping person as secretor or nonsecretor
XX Disclosure; Column 45-50; 55pp; English.
XX This sequence is the human Sec2 protein of the invention. The DNA
CC encodes a alpha(1,2) fucosyltransferase and is the Secretor
CC alpha(1,2) fucosyltransferase locus, that cross hybridises with the
CC H blood group alpha(1,2) fucosyltransferase gene. The DNA is useful for
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC individual as a secretor or nonsecretor as it is known that nonsecretors
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC the position corresponding to amino acid 143.
XX Sequence 343 AA;
Query Match 73.1%; Score 1386.5; DB 19; Length 343;
Best Local Similarity 79.1%; Pred. No. 7.3e-143;
Matches 256; Conservative 26; Mismatches 31; Indels 11; Gaps 2;
QY 1 LQORIVKLQPLSEKELPMTTOMSSGNTESPMPRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 28 VQQLAKTIQAN--WELPV-----QIPVLASTKALGPSQLRGMTINAIKRLGNQM 76
QY 61 GEYATLALARMNGRLAFIPASMHNALAFIPRISLPVLHSDTAKKIPWNYHLNDWMEER 120
Db 77 GEYATLALAKMNGRPAFIPAQMHSTLAFIPRITLPVLHSTASRIPWNYHLNDWMEER 136
QY 121 YRHIPGHVFTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQPSTFV 180
Db 137 YRHIPGEYVRLTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSRPSTFV 196
QY 161 GVHVRGDYVHVMNVKGVYVADRGYLEKALDMFRARYSPFVVTSGNAWCNENAS 240
Db 197 GVHVRGDYVHVMNVKGVYVADRGYLEKALDMFRARYSPFVVTSGNAWCNENAS 256

Db 59 GTAMGNASSCPQH-PASLSGTWVYPNGRFGNQMGQVATLLALAQNGRRRAFILPAMH 117
 QY 85 NALAPFRLSLVLDSDAKKTPWQNYHLNDWMEERYRHIPGHEVFTGYPGCSWTFYHHL 144
 Db 118 AALAPVFRITLVLAPEDVSRTPWRELQHDWMESEYADLRDPFLKLSGFPSCWTFYHHL 177
 QY 145 RPEILKEFTLHDHVRREEAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHVMNVKGVVA 202
 Db 178 REQIRREFTLHDHVRREEAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHVMNVKGVVA 237
 QY 203 DRGYLEKALDMFRARYSSPVFVVTNSGMAWCRENINASRGDVVVFAGNIEGSPAKDFALL 262
 Db 238 DSAYLRQAMDWFRARHEAPVFFVVTNSGMECKENITDSQDVTFAAGDQGEATPWKDFALL 297
 QY 263 TCQNTHTIMTIGTGFVAAAYLAGGDTIYLANFTLPDSEFLKIFKPEAAFLPEWVGIPADLS 322
 Db 298 TCQNTHTIMTIGTGFVAAAYLAGGDTIYLANFTLPDSEFLKIFKPEAAFLPEWVGIPADLS 357
 QY 323 PLLKALTP 330
 Db 358 PLWTLAKP 365

RESULT 10
 ID AAR80154 standard; Protein; 365 AA.
 AC AAR80154;
 XX
 XX 25-JUN-1996 (first entry)
 DT
 DE GDP-L-fucose-beta-D-galactosidase alpha-1,2-fucosyltransferase.
 KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase; Fuc-T;
 KW synthesis; 2'-fucosyllactose; oligosaccharide; human milk;
 KW non-human transgenic mammal; secondary gene product.
 XX
 OS Homo sapiens.
 XX
 XX W09524494-A1.
 XX
 XX 14-SEP-1995.
 PD
 XX 24-JAN-1995; 95WO-US00926.
 XX
 XX 09-MAR-1994; 94US-0209122.
 XX
 XX (ABBO) ABBOTT LAB.
 XX
 XX Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
 PI Pierce JM, Prieto PA, Smith DF;
 XX
 XX WPI: 1995-336739/43.
 DR N-PSDB; AAO98461.
 XX
 XX Prodn. of human:ised milk by non-human transgenic mammal - by
 PT inserting heterologous gene encoding human catalytic entity which
 PT produces oligo:saccharide(s) and glyco:conjugate(s) into mammal's
 PT genome
 XX
 XX Example 1; Page 62-64; 83pp; English.
 XX
 XX GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyl-transferase (Fuc-T)
 CC is encoded by cDNA (AAO98461) isolated from a human epidermal carcinoma
 CC cell line. The enzyme is responsible for the synthesis of
 CC 2'-fucosyllactose, one of the oligosaccharides in human milk. Other
 CC products of Fuc-T include glycoproteins contg. beta-linked terminal
 CC galactose residues which can be fucosylated by Fuc-T. This DNA can
 CC be used to produce non-human transgenic mammals able to produce
 CC secondary gene products, e.g. oligosaccharides, in their milk. The
 CC transgenic mammals milk biochemically resembles human milk. This
 CC humanised milk can be used in the prepn. of an enteral nutritional
 CC product useful in the nutritive maintenance of an animal.

XX Sequence 365 AA;
 Query Match 55.5%; Score 1051.5; DB 16; Length 365;
 Best Local Similarity 63.3%; Pred. No. 4e-106;
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
 QY 25 GNTGSEPMERDSEHGNGELRGMTINSIGRLGNQMGQVATLLALAQNGRRRAFILPAMH 84
 Db 59 GTAMGNASSCPQH-PASLSGTWVYPNGRFGNQMGQVATLLALAQNGRRRAFILPAMH 117
 QY 85 NALAPFRLSLVLDSDAKKTPWQNYHLNDWMEERYRHIPGHEVFTGYPGCSWTFYHHL 144
 Db 118 AALAPVFRITLVLAPEDVSRTPWRELQHDWMESEYADLRDPFLKLSGFPSCWTFYHHL 177
 QY 145 RPEILKEFTLHDHVRREEAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHVMNVKGVVA 202
 Db 178 REQIRREFTLHDHVRREEAQAFRLGRV--NGSQPSTFVGHVHVRGDDYVHVMNVKGVVA 237
 QY 203 DRGYLEKALDMFRARYSSPVFVVTNSGMAWCRENINASRGDVVVFAGNIEGSPAKDFALL 262
 Db 238 DSAYLRQAMDWFRARHEAPVFFVVTNSGMECKENITDSQDVTFAAGDQGEATPWKDFALL 297
 QY 263 TCQNTHTIMTIGTGFVAAAYLAGGDTIYLANFTLPDSEFLKIFKPEAAFLPEWVGIPADLS 322
 Db 298 TCQNTHTIMTIGTGFVAAAYLAGGDTIYLANFTLPDSEFLKIFKPEAAFLPEWVGIPADLS 357
 QY 323 PLLKALTP 330
 Db 358 PLWTLAKP 365

RESULT 11
 AAR70422
 ID AAR70422 standard; Protein; 365 AA.
 XX
 XX AAR70422;
 AC
 XX 27-FEB-1996 (first entry)
 DT
 DE 2-Alpha-fucosyltransferase.
 XX
 KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
 KW glycosyltransferase; glycosylation; oligosaccharide; glycoprotein;
 KW glycolipid; transgenic animal; cattle; milk.
 XX
 OS Homo sapiens.
 XX
 XX W09524495-A1.
 XX
 XX 14-SEP-1995.
 PD
 XX 24-JAN-1995; 95WO-US00967.
 XX
 XX 09-MAR-1994; 94US-0208889.
 XX
 XX (ABBO) ABBOTT LAB.
 XX
 XX Cummings RD, Kopchik JJ, Moreman KW, Mukerji P;
 PI Pierce JM, Prieto PA, Smith DF;
 XX
 XX WPI: 1995-328284/42.
 DR N-PSDB; AAT01083.
 XX
 XX New transgenic non-human mammal milk prods - contg. heterologous
 PT components produced as secondary gene prods. of an heterologous gene
 XX
 XX Example 1; Page 62-64; 83pp; English.
 PS
 XX 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone
 CC (AAT01083) isolated from a human epidermal carcinoma A431 library.
 CC The enzyme can be expressed in the milk of a transgenic mammal,
 CC esp. cow. This allows large-scale prodn. of oligosaccharides and

CC glycosylated proteins and lipids in the milk.

XX Sequence 365 AA;

SQ Query Match 55.5%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRDRSEHGNGELRCMFTINSIGRLGNOMGEYATLAFALARMNGRLAFIPASMH 84
DB 59 GTAMPNASSCCPQH-PASLSGTWTVYVNGRFGNMGQYATLALALQNLGRRAFIIPAMH 117
QY 85 NALAPFRISLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHVRFYGYPCSWTFYHHL 144
DB 118 AALAPVFRITLPLVLAPEVDSRTFWRQLQHDWMSBYADLRDPFLKSGFPCSWTFYHHL 177
QY 145 REELIKEFTLHDHREEAQAFGLRLV--NGSQPSTFVGVHVRGDDYVHVMNVKGVVA 202
DB 178 REQIRREFTLHDHREEAQSVLGQLRGTRGDRPRTFVGVHVRGDDYVHVMNVKGVVG 237
QY 203 DRGYLEKALDMFRARYSSPVVYTSNGMAWCRENINASRGDVVFAGNGTEGSPAKDFALL 262
DB 238 DSAYLQAMDWRARHEAPVYVYTSNGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 263 TCNHTIMTIGTFGIWAAYLAGDITVYLANITLPSFPLKVKPEAAFLPEWVGIPADLS 322
DB 298 TCNHTIMTIGTFGIWAAYLAGDITVYLANITLPSFPLKVKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365

RESULT 12

AAR70421
ID AAR70421 standard; Protein: 365 AA.

XX AAR70421;

DI 27-FEB-1996 (first entry)

XX 2-Alpha-fucosyltransferase.

XX GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
KW transgenic animal; cattle; glycosylation; milk.

XX Homo sapiens.

XX WO9524488-A1.

XX 14-SEP-1995.

XX 24-JAN-1995; 95WO-US01147.

XX 09-MAR-1994; 94US-0209132.

XX (ABSO) ABBOTT LAB.

XX Cummings RD, Kopechik JJ, Moremen KW, Mukerji P;

XX Pierce JW, Prieto PA, Smith DF;

XX WPI; 1995-328279/42.

XX N-PSDB; AAT01082.

XX Transgenic animal expressing heterologous catalyst - used in
PT metabolite prodn esp. glycosyltransferase for prodn. of
PT oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
PT foods, pharmaceuticals, etc.

XX Example 1; Page 62-64; 84pp; English.

XX Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone

CC (AAT01082) isolated from a human epidermal carcinoma cell line A431
CC cDNA library. The enzyme can be expressed in the milk of a
CC transgenic mammal, esp. cow. This allows large-scale prodn. of
CC oligosaccharides or glycosylated proteins and lipids in the milk.

XX Sequence 365 AA;

SQ Query Match 55.5%; Score 1051.5; DB 16; Length 365;
Best Local Similarity 63.3%; Pred. No. 4e-106;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSEPMRDRSEHGNGELRCMFTINSIGRLGNOMGEYATLAFALARMNGRLAFIPASMH 84
DB 59 GTAMPNASSCCPQH-PASLSGTWTVYVNGRFGNMGQYATLALALQNLGRRAFIIPAMH 117
QY 85 NALAPFRISLPLVLSHTAKKIPQWNYHLNDWMEERYRHIPGHVRFYGYPCSWTFYHHL 144
DB 118 AALAPVFRITLPLVLAPEVDSRTFWRQLQHDWMSBYADLRDPFLKSGFPCSWTFYHHL 177
QY 145 REELIKEFTLHDHREEAQAFGLRLV--NGSQPSTFVGVHVRGDDYVHVMNVKGVVA 202
DB 178 REQIRREFTLHDHREEAQSVLGQLRGTRGDRPRTFVGVHVRGDDYVHVMNVKGVVG 237
QY 203 DRGYLEKALDMFRARYSSPVVYTSNGMAWCRENINASRGDVVFAGNGTEGSPAKDFALL 262
DB 238 DSAYLQAMDWRARHEAPVYVYTSNGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
QY 263 TCNHTIMTIGTFGIWAAYLAGDITVYLANITLPSFPLKVKPEAAFLPEWVGIPADLS 322
DB 298 TCNHTIMTIGTFGIWAAYLAGDITVYLANITLPSFPLKVKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365

RESULT 13

AAR90572

ID AAR90572 standard; Protein: 365 AA.

XX AAR90572;

XX 08-APR-1996 (first entry)

XX Human H-transferase.

XX H-transferase; xenograft hyperacute rejection; transplantation;
KW glycosyltransferase; galactose alpha(1,3) galactose.

XX Homo sapiens.

XX WO9534202-A1.

XX 21-DEC-1995.

XX 14-JUN-1995; 95WO-US07554.

XX 21-JUL-1994; 94US-0278282.

XX 15-JUN-1994; 94US-0260201.

XX (ALEX-) ALEXION PHARM INC.

XX (AUST-) AUSTIN RES INST.

XX Fodor WL, McKenzie IFC, Rother RP, Sandrin MS; Squinto SP;

XX WPI; 1996-049326/05.

XX N-PSDB; AAT12238.

XX Redn. of rejection of xenogeneic cells following transplantation
PT by introducing a vector expressing fucosyltransferase into the
PT cells

XX Example 1; Page 45-47; 69pp; English.

XX	The human H-transferase (AAR90572) product of a cDNA clone (AAT12238);
CC	prep'd. from human epidermoid carcinoma cells (ATCC CRL 155 A-431)
CC	can be expressed in xenogeneic organs, tissues and cells using
CC	a vector such as pAPEX-1 (AAT12239). This results in decreased
CC	expression of the non-human antigen galactose alpha(1,3) galactose
CC	on the surface of the organs etc. so that hyperacute rejection is
CC	reduced upon transplantation to humans.
XX	
SQ	Sequence 365 AA:
	Query Match 55.5%; Score 1051.5; DB 17; Length 365;
	Best Local Similarity 63.3%, Pred. No. 4e-106;
	Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps
QY	25 GNTPSEMRERDSQHNGELGRGNTINSIGRLGNOMGEYATILFALARNMGRLAFTASNMH 84
DB	: :
59 GTANGPNASSCPQH--PASLSGTWTVPVNGRFGNQGVATLLALAQLNRRRAFTLPAMH 117	
QY	85 NALAPTRISLPLVLHSDTAKKIPWQNYHLNDWMEEYRHIPCHGFVRFTGYPCSWTFYHHL 144
DB	: :
118 AALPAVFRTITPLVAPEVDSPTRPWRLEQLHDNMSEBYADLRDPFEKLSGFPCSMTFFHRL 177	
QY	145 RPILLKEFTLHDHVREDAFAFLGLRV--NGSQPTFCVGHVRRGDYVHMVNWKGVYA 202
DB	: :
178 REQIRREFTLHDHREAEOSVGLQLKLGRTGDRPTFGVHVHVRGDYLQVMQRKWGVWG 237	
QY	203 DRGYLEKALDMFRARYSSPVFYVTSNGMAWCRENINASRGDVVVFNAGNTEGSPAKDFALL 262
DB	:
238 DSAYLCAMDWFARIEAPVFVYVTSNGMEWCKENIDTSQDVTTFAGDGQEAETPKDFALL 297	
QY	263 TCQNHTIMTTGTGGIWAAYLAGDDTIYLANLYLPDSPFFLKVKPEAAFTPEWVGIPADLS 322
DB	
298 TCQNHTIMTTGTGFWAAYLAGDVTYLANFTLPDSEFLKIKPEAAFTPEWVGINADLS 357	
QY	323 PLLKALT P 330
DB	
358 PLWTLAK P 365	
RESULT 14	
AARW23805	
ID	AARW23805 standard; Protein; 365 AA.
AC	AARW23805;
AC	
DT	15-SEP-1997 (first entry)
XX	
DE	Human alpha 1,2 fucosyltransferase (EC 2.4.1.69).
XX	
KW	Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;
KW	alpha 1,2 F _T ; transgene; transgenic mouse; animal model;
KW	intestinal adhesion; Helicobacter pylori infection; stomach;
KW	small intestine; gut; epithelial cell; surface receptor;
KW	carbohydrate antigen; gastritis; peptic ulcer; neoplasia;
KW	gastric adenocarcinoma; Lewis antigen; fucosylation.
OS	Homo sapiens.
XX	
PN	US5625124-A.
XX	
PD	29-APR-1997.
XX	
PF	11-JUL-1994; 94US-0273411.
XX	
PR	11-JUL-1994; 94US-0273411.
XX	
PA	(UNIW) UNIV WASHINGTON.
XX	
PI	Falk P, Gordon JI;
XX	
DR	WPI; 1997-258275/23.
NR	N-PSDB; AAT76768.

```
XX (UNMI ) UNIV MICHIGAN.
PA
XX Legault DJ, Lowe JB;
XX WPI; 1997-192897/17.
XX N-PSDB; AAT61677.
XX
XX New recombinant fucosyltransferase proteins - useful for modifying
XX cell surface oligosaccharide structures
XX
XX Example 1; Page 280-281; 329pp; English.
XX
XX Human GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase
XX (AAW13640) is an enzyme involved in the expression of type I and II
XX blood group H structures. Its amino acid sequence was deduced
XX from an isolated DNA (AAT61677) derived from human A431 cells.
XX Expression of the fucosyltransferase animal cell lines provides
XX specific capabilities with respect to post-translational
XX modification of the oligosaccharides of expressed proteins or
XX lipids. Specific applns. of the enzyme include enzymatic
XX fucosylation of chain-terminating galactose residues on
XX lactosamine or neolacto type beta-D-galactoside to alpha-2-
XX fucose residues. The enzyme can also be used to raise antibodies
XX as diagnostic reagents and to screen cpds. for fucosyltransferase
XX inhibitor activity.
XX
XX SQ Sequence 365 AA;
XX
XX Query Match 55.5%; Score 1051.5; DB 18; Length 365;
XX Best Local Similarity 63.3%; Pred. No. 4e-106;
XX Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
XX
XX QY 25 GNTSPENRRSEQHNGELRCQFTINSIGRLGNOMGEYAILFALARMNGRLAFIPASMH 84
XX DB 59 GTAMGNASSQCPQH-PASLSGTWTVPYNGRFGNQMGQYATLLAQLNGRAFTLPAMH 117
XX QY 85 NALAPFRLSLPVLHSDTAKKIPWQNYHLNDWMEERYRHIPGHVFRFTGYPCSWTFYHHL 144
XX DB 118 AALAPVFRITLPLVLAPEVDSRTPWRELQLHDWMSSEYADLRDPFLKLSGFCPSWTFYHHL 177
XX QY 145 RPEILKEETLHDHVREREAQAFRLRLV--NGSQPSTFVGHVHVRGDIYHVMPNVMKGVYA 202
XX DB 178 REQIRREFTLHDHVREREAQSVLQLRLGRTGDRPTFGVGVHVRGDIYLVQVMPQRWKGVYG 237
XX QY 203 DRGYLEKALDMFRARYSSPVFVVTNSGMAWCARENINASEGDVVFAGNGIEGSPAKDFALL 262
XX DB 238 DSAYLRQAMDFRAHEAPVFVVTNSGMEWCKENIDTSQGDVTFAGDGOEATPWKDFALL 297
XX QY 263 TCQNTHTIMTIGTFGIWAAYLAGGDIYLANVYLPDPPFLKFKPEAAFLPEWVGIPADLS 322
XX DB 298 TCQNTHTIMTIGTFGFWAAYLAGGDIYLANVYLPDSEFLKFKPEAAFLPEWVGINADLS 357
XX QY 323 PLLKALTP 330
XX DB 358 PLWTLAKP 365
```

Search completed: May 27, 2003, 15:06:16
Job time : 45.7707 secs

us-10-040-863-10.ra1

Wed May 28 09:21:51 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 ; Search time 14.5136 Seconds
(without alignments)
715.625 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQORIVKLPSEKELPMTT.....RSHFHLXAKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cn2.6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cn2.6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cn2.6/ptodata/1/1aa/6A-COMB.pep:*
- 4: /cn2.6/ptodata/1/1aa/6B-COMB.pep:*
- 5: /cn2.6/ptodata/1/1aa/6C-COMB.pep:*
- 6: /cn2.6/ptodata/1/1aa/6D-COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1896	100.0	333	4	US-09-298-886-10
2	1896	100.0	380	4	US-09-298-886-5
3	1430	75.4	340	4	US-09-254-077A-6
4	1430	75.4	340	4	US-09-254-077A-7
5	1386.5	73.1	343	1	US-08-395-800A-8
6	1386.5	73.1	343	1	US-09-254-077A-8
7	1376	72.6	344	4	US-09-298-886-11
8	1375.5	72.5	347	4	US-09-254-077A-9
9	1080.5	57.0	373	4	US-09-254-077A-12
10	1051.5	55.5	333	5	PCT-US91-00899-11
11	1051.5	55.5	365	1	US-07-914-281-6
12	1031.5	55.5	365	1	US-08-393-246-6
13	1051.5	55.5	365	1	US-08-273-411-1
14	1051.5	55.5	365	1	US-08-525-058A-6
15	1051.5	55.5	365	1	US-08-395-800A-6
16	1051.5	55.5	365	1	US-08-395-800A-10
17	1051.5	55.5	365	2	US-08-696-731-6
18	1051.5	55.5	365	4	US-09-042-531-6
19	1051.5	55.5	365	4	US-09-254-077A-11
20	1051.5	55.5	365	5	PCT-US91-00899-12
21	1046.5	55.2	365	4	US-09-151-592-2
22	1046.5	55.2	365	4	US-09-254-077A-10
23	691	36.4	222	1	US-08-395-800A-2
24	439	23.2	102	1	US-08-393-800A-3
25	141	7.4	300	4	US-09-433-598-2
26	97	5.1	183	1	US-09-167-035-33
27	97	5.1	183	1	US-08-208-887A-33

28	97	5.1	183	2	US-08-539-005-33	Sequence 33, Appl
29	97	5.1	183	4	US-09-280-598-35	Sequence 35, Appl
30	96.5	5.1	355	4	US-09-390-131-5	Sequence 5, Appl
31	91.5	4.8	801	1	US-07-906-349A-6	Sequence 6, Appl
32	91	4.8	442	1	US-08-220-151-22	Sequence 22, Appl
33	91	4.8	442	1	US-08-413-118-22	Sequence 22, Appl
34	91	4.8	442	3	US-08-473-446-22	Sequence 22, Appl
35	91	4.8	442	6	5470718-3	Patent No. 5470718
36	87	4.6	217	1	US-08-167-035-6	Sequence 6, Appl
37	87	4.6	217	1	US-08-208-887A-6	Sequence 6, Appl
38	87	4.6	217	2	US-08-539-005-6	Sequence 6, Appl
39	87	4.6	217	2	US-08-815-176-3	Sequence 3, Appl
40	87	4.6	217	2	US-08-815-176-4	Sequence 4, Appl
41	87	4.6	217	4	US-08-664-962B-6	Sequence 6, Appl
42	87	4.6	217	4	US-09-311-743-6	Sequence 6, Appl
43	87	4.6	217	4	US-09-280-598-6	Sequence 6, Appl
44	87	4.6	217	4	US-09-197-344-3	Sequence 3, Appl
45	87	4.6	217	4	US-09-197-344-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-298-886-10

; Sequence 10, Application US/09298886

; Patent No. 6329170

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/298,886

; CURRENT FILING DATE: 1999-04-26

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 10

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; ORGANISM: Rattus norvegicus

US-09-298-886-10

Query Match 100.0%; Score 1896; DB 4; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.6e-210;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LQORIVKLPSEKELPMTTOMSSGNTESPEMRDSEQHGNGELRGMTTINSIGRLGNQM	60
DB	1	LQORIVKLPSEKELPMTTOMSSGNTESPEMRDSEQHGNGELRGMTTINSIGRLGNQM	60
QY	61	GEYATFLALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER	120
DB	61	GEYATFLALARNGLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEER	120
QY	121	YHHPGHVFRFTGYPSCSTFVYHHLRPEILKEFTLHDHVREREAQAFRLGVNGSQPSTFV	180
DB	121	YHHPGHVFRFTGYPSCSTFVYHHLRPEILKEFTLHDHVREREAQAFRLGVNGSQPSTFV	180
QY	181	GVHVRGDIYHVMNPVWKGVDRCYLEKALDMFRARYSSPVFVVTSSNGMWCNENINAS	240
DB	181	GVHVRGDIYHVMNPVWKGVDRCYLEKALDMFRARYSSPVFVVTSSNGMWCNENINAS	240
QY	241	RGDVVFAGNIEGSPAKDPALLTQCNHTTMTGTGFWAAYLAGDITIYLYNTLPDPSF	300
DB	241	RGDVVFAGNIEGSPAKDPALLTQCNHTTMTGTGFWAAYLAGDITIYLYNTLPDPSF	300
QY	301	LKVFKPEAAFLPEWVGIPADLPALLKALTPACPRSHFHLKAGVTCYVAGRAF	353
DB	301	LKVFKPEAAFLPEWVGIPADLPALLKALTPACPRSHFHLKAGVTCYVAGRAF	353

RESULT 2

US-09-298-886-8
 ; Sequence 8, Application US/09298886
 ; Patent No. 6329170

; GENERAL INFORMATION:
 ; APPLICANT: ERIC H. HOLMES ET AL.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
 ; TITLE OF INVENTION: GMI-SPECIFIC ALPHA-2 FUCOSYLTRANSFERASE AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 8511-029
 ; CURRENT APPLICATION NUMBER: US/09/298,886
 ; CURRENT FILING DATE: 1999-04-26
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-298-886-8

Query Match 100.0%; Score 1896; DB 4; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.8e-210;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 88 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 147
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 148 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 207
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 208 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 257
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 300
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 268 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 327
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 301 LKVFKEPAAFLPEWVGIPADLSPLK 326
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 328 LKVFKEPAAFLPEWVGIPADLSPLK 339
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 3

US-09-254-077A-6
 ; Sequence 6, Application US/09254077A
 ; Patent No. 6399758

; GENERAL INFORMATION:
 ; APPLICANT: SANDRIN, MAURO S.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; FILE REFERENCE: 30562.5USWO
 ; CURRENT APPLICATION NUMBER: US/09/254,077A
 ; CURRENT FILING DATE: 1999-06-11
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00540
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: PO 1823
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-6

Query Match

Best Local Similarity

75.4%; Score 1430; DB 4; Length 340;

81.9%; Pred. No. 1.1e-156;

Matches 267; Conservative

26; Mismatches 19; Indels

14; Gaps

3;

```

QY 1 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 73
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 133
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 134 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 193
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 194 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 253
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 300
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 254 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 313
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 301 LKVFKEPAAFLPEWVGIPADLSPLK 326
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 314 LKVFKEPAAFLPEWVGIPADLSPLK 339
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 4

US-09-254-077A-7

; Sequence 7, Application US/09254077A
 ; Patent No. 6399758

; GENERAL INFORMATION:
 ; APPLICANT: SANDRIN, MAURO S.

; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; FILE REFERENCE: 30562.5USWO
 ; CURRENT APPLICATION NUMBER: US/09/254,077A
 ; CURRENT FILING DATE: 1999-06-11
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00540
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: PO 1823
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-7

Query Match

Best Local Similarity

75.4%; Score 1430; DB 4; Length 340;

81.9%; Pred. No. 1.1e-156;

Matches 267; Conservative

26; Mismatches 19; Indels

14; Gaps

3;

```

QY 1 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 LQORIYKLOPLSEKELPMTTQSSNGTSEPMRRDSEQHNGELRGMTINSTIGRLGNOM 73
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 GEYATLALARMNGRLAFIPASHMHNALAFIRISLPVLHSDTAKKIPWNYHLNDWMEER 133
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 134 YRHIPGHEVRFYGYPCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSTFV 193
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 194 GVHVRGDIYVHVPNTWKGVADRGYLEKALDMFRARYSPFVVTSSNGMACKREINAS 253
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 300
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 254 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTGTGFIWAAYLAGGDTIYLANITLDPSPF 313
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Suite 400
; City: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00899
; FILING DATE: 19910214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye Ph.D., Jean-Paul
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-5940
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Blood
; PCT-US91-00899-11

```

```

Query Match      55.5%; Score 1051.5; DB 5; Length 333;
Best Local Similarity 63.3%; Pred. No. 6.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 NTESPEMRDRSEQHNGELRGMTINSIGRLGNMGVATLALARMNGRLAFIPASMH 84
Db 27 GTAMGPNASSCPQH-PASLSGTWTVPNGRGNMGVATLALARMNGRLAFIPASMH 85
QY 85 NALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHFVFTGYPGCSWTFYHHL 144
Db 86 AALAPVFRITLFLVLAPEVDSRTPWRELQLDHWMSEYADLRDPELKLSPGCSWTFYHHL 145
QY 145 RPEILKEFTLHDHVRREAAQFLRLRV--NGSQPSTFVGVHVRGDIYVHVPNQWKGVA 202
Db 146 REQIRREFTLHDHVRREAAQSVLGRLGRTGDRPTFVGVHVRGDIYVHVPNQWKGVA 205
QY 203 DRGYLEALDMFRARYSSPVFVTSNGMWCNENINASRGDVFAGNGIEGSPAKDFALL 262
Db 206 DSAYLRQAMDWFARHEAPVFTVTSNGMWCNENINASRGDVFAGNGIEGSPAKDFALL 265
QY 263 TCQNHHTMTTGTGFWAAAYLAGGDTIYLANFTLPDSFPLKVKPEAAFLPEWVGIPADLS 322
Db 266 TCQNHHTMTTGTGFWAAAYLAGGDTIYLANFTLPDSFPLKVKPEAAFLPEWVGIPADLS 325
QY 323 PLLKALTP 330
Db 326 PLWTLAKP 333

```

```

RESULT 11
US-07-914-281-6
; Sequence 6, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.

```

```

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; City: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-914-281-6

```

```

Query Match      55.5%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 7.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 NTESPEMRDRSEQHNGELRGMTINSIGRLGNMGVATLALARMNGRLAFIPASMH 84
Db 59 GTAMGPNASSCPQH-PASLSGTWTVPNGRGNMGVATLALARMNGRLAFIPASMH 117
QY 85 NALAPIRISLPVLHSDTAKKIPWQNYHLNDWMEERYHIFGHFVFTGYPGCSWTFYHHL 144
Db 118 AALAPVFRITLFLVLAPEVDSRTPWRELQLDHWMSEYADLRDPELKLSPGCSWTFYHHL 177
QY 145 RPEILKEFTLHDHVRREAAQFLRLRV--NGSQPSTFVGVHVRGDIYVHVPNQWKGVA 202
Db 178 REQIRREFTLHDHVRREAAQSVLGRLGRTGDRPTFVGVHVRGDIYVHVPNQWKGVA 237
QY 203 DRGYLEALDMFRARYSSPVFVTSNGMWCNENINASRGDVFAGNGIEGSPAKDFALL 262
Db 238 DSAYLRQAMDWFARHEAPVFTVTSNGMWCNENINASRGDVFAGNGIEGSPAKDFALL 297
QY 263 TCQNHHTMTTGTGFWAAAYLAGGDTIYLANFTLPDSFPLKVKPEAAFLPEWVGIPADLS 322
Db 298 TCQNHHTMTTGTGFWAAAYLAGGDTIYLANFTLPDSFPLKVKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
Db 358 PLWTLAKP 365

```

```

RESULT 12
US-08-393-246-6
; Sequence 6, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.

```



```

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,246
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-393-246-6

```

```

Query Match 55.5%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 7.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSPENRRDSEOHNGELRCGMFTINSIGRLGNQMGEXATLFLALRMNGRLAIPASMH 84
Db 59 GTAMGPNASSSCPOH-PASLSGTTWVPNGRFGNMGQYATLLAALQNLGRRAPILPAMH 117
QY 95 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMBERYRHIPGHVFTGPGCSWTFYHHL 144
Db 118 AALAPVFRITLPLVAPEVDSRTPWRELQHDWSEEXADLRDPFLKLSGPGCSWTFHHL 177
QY 145 RPEILKEFTLHDHVREREAQAFLGLRV--NGSQPSTFVGHVHRGDYVHVMPNWKGVVA 202
Db 178 REQIRREFTLHDHVREREAQAFLGLRV--NGSQPSTFVGHVHRGDYVHVMPNWKGVVA 237
QY 203 DRGYLEKALDMFRARYSSPVFTVTSNMGWCKENINASRGDVPVFNAGNIGSPAKDFALL 262
Db 238 DSAYLQAMDFRAREAPVFTVTSNMGWCKENINASRGDVPVFNAGNIGSPAKDFALL 297
QY 263 TCQNTHTMTTGTGFWAAYLAGDVTIYANLTPDPSFLKVKPEAFLPEWVGIPADLS 322
Db 298 TCQNTHTMTTGTGFWAAYLAGDVTIYANLTPDPSFLKVKPEAFLPEWVGIPADLS 357
QY 323 PLKALTP 330
Db 358 PLWTAKP 365

```

RESULT 13

```

US-08-273-411-1
; Sequence 1, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..365
; OTHER INFORMATION: /note="GDP-L-fucose:beta-D-Galactoside-2-alpha-L-fucosyl
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 365
; US-08-273-411-1

```

```

Query Match 55.5%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 7.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTSPENRRDSEOHNGELRCGMFTINSIGRLGNQMGEXATLFLALRMNGRLAIPASMH 84
Db 59 GTAMGPNASSSCPOH-PASLSGTTWVPNGRFGNMGQYATLLAALQNLGRRAPILPAMH 117
QY 95 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMBERYRHIPGHVFTGPGCSWTFYHHL 144
Db 118 AALAPVFRITLPLVAPEVDSRTPWRELQHDWSEEXADLRDPFLKLSGPGCSWTFHHL 177
QY 145 RPEILKEFTLHDHVREREAQAFLGLRV--NGSQPSTFVGHVHRGDYVHVMPNWKGVVA 202
Db 178 REQIRREFTLHDHVREREAQAFLGLRV--NGSQPSTFVGHVHRGDYVHVMPNWKGVVA 237
QY 203 DRGYLEKALDMFRARYSSPVFTVTSNMGWCKENINASRGDVPVFNAGNIGSPAKDFALL 262

```

Db 238 DSAYLRQAMDFRARHEAPVFFVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALL 297
Qy 263 TOCNHTIMTIGTGWAAYLGGDDTIYLANFTLPDSPLKVKFKPEAAFLPEWVGIPADLS 322
Db 298 TOCNHTIMTIGTGWAAYLGGDDTIYLANFTLPDSPLKVKFKPEAAFLPEWVGINADLS 357
Qy 323 PLLKALTP 330
Db 358 PLWTLAKP 365
RESULT 14
US-08-525-058A-6
; Sequence 6, Application US/08525058A
; Patent No. 5770420
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS.
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,058A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-058A-6
Query Match 55.5%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 7.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
Qy 25 GNTSPENRRDSEOHNGELGMFTINSIGRLGNOMGEYATILFALARMNGRIAFIPASMH 84
Db 59 GTAMGNASSCPQH-PASLSGTWTVPYNGRFGNOMGYATLLAQLNGRAFLTPAMH 117
Qy 85 NALAPIFRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHFVRRFGYPCSWTFYHHL 144
Db 118 AALAPVFRITLPVLAPEVDSRTPRELQHDWMSSEYADLRDPFLKLSGFPSCWTFYHHL 177
Qy 145 RPEILKETLHDHVREREAQAFLGRUV--NGSQPTFFGVHVRGDYVHVMPNWKGVVA 202
Db 178 RQIRREFTLHDHLEEAQAQSVLGQLRLGRTGDRPRTFVGHVHVRGDYLVQVMPQRWKGVVG 237
Qy 203 DRGYLEKALDMFRARYSSPVFVVTNSGMWCRENINASRGDVFVFNAGNTEGSPAKDFALL 262
Db 238 DSAYLRQAMDFRARHEAPVFFVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALL 297

Qy 263 TOCNHTIMTIGTGWAAYLGGDDTIYLANFTLPDSPLKVKFKPEAAFLPEWVGIPADLS 322
Db 298 TOCNHTIMTIGTGWAAYLGGDDTIYLANFTLPDSPLKVKFKPEAAFLPEWVGINADLS 357
Qy 323 PLLKALTP 330
Db 358 PLWTLAKP 365
RESULT 15
US-08-395-800A-6
; Sequence 6, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J.
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-395-800A-6
Query Match 55.5%; Score 1051.5; DB 1; Length 365;
Best Local Similarity 63.3%; Pred. No. 7.2e-113;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
Qy 25 GNTSPENRRDSEOHNGELGMFTINSIGRLGNOMGEYATILFALARMNGRIAFIPASMH 84
Db 59 GTAMGNASSCPQH-PASLSGTWTVPYNGRFGNOMGYATLLAQLNGRAFLTPAMH 117
Qy 85 NALAPIFRISLPVLHSDTAKKIPWNYHLNDWMEERYRHIPGHFVRRFGYPCSWTFYHHL 144
Db 118 AALAPVFRITLPVLAPEVDSRTPRELQHDWMSSEYADLRDPFLKLSGFPSCWTFYHHL 177
Qy 145 RPEILKETLHDHVREREAQAFLGRUV--NGSQPTFFGVHVRGDYVHVMPNWKGVVA 202
Db 178 RQIRREFTLHDHLEEAQAQSVLGQLRLGRTGDRPRTFVGHVHVRGDYLVQVMPQRWKGVVG 237
Qy 203 DRGYLEKALDMFRARYSSPVFVVTNSGMWCRENINASRGDVFVFNAGNTEGSPAKDFALL 262
Db 238 DSAYLRQAMDFRARHEAPVFFVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALL 297

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:08:29 ; Search time 15.2516 Seconds
(without alignments)
2295.367 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKLPISKEPLMTT.....RSHEFLKAKGYCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1896	100.0	353	10	US-09-999-672-10
2	1896	100.0	353	12	US-10-040-863-10
3	1896	100.0	380	10	US-09-999-672-8
4	1896	100.0	380	12	US-10-040-863-8
5	1430	75.4	340	10	US-09-051-034A-2
6	1376	72.6	344	10	US-09-999-672-11
7	1376	72.6	344	12	US-10-040-863-11
8	1051.5	55.5	365	9	US-10-105-963-10
9	1051.5	55.5	365	10	US-09-863-475A-6
10	1046.5	55.2	365	10	US-09-051-034A-4
11	1044.5	55.1	365	10	US-09-844-268-13
12	1044.5	55.1	365	10	US-09-844-705-13
13	141	7.4	300	10	US-09-848-838-2
14	99	5.2	436	9	US-10-214-524-25
15	97.5	5.1	431	9	US-09-479-614-14
16	97.5	5.1	436	9	US-09-479-614-2
17	97.5	5.1	436	9	US-09-479-614-29
18	89	4.7	1031	9	US-09-764-868-647
19	89	4.7	2057	9	US-09-815-379-10

20	89	4.7	2058	9	US-09-815-379-17	Sequence 17, Appl
21	87	4.6	217	10	US-09-765-298A-6	Sequence 6, Appl
22	85.5	4.5	738	9	US-10-235-521-1	Sequence 1, Appl
23	84.5	4.5	344	9	US-09-895-913A-56	Sequence 56, Appl
24	84.5	4.5	344	10	US-09-815-242-11366	Sequence 11366, A
25	83.5	4.4	747	9	US-10-028-072-212	Sequence 212, App
26	83.5	4.4	747	9	US-10-121-049-212	Sequence 212, App
27	83.5	4.4	747	9	US-10-123-904-212	Sequence 212, App
28	83.5	4.4	747	9	US-10-140-470-212	Sequence 212, App
29	83.5	4.4	747	9	US-10-175-746-212	Sequence 212, App
30	83.5	4.4	747	9	US-10-176-918-212	Sequence 212, App
31	83.5	4.4	747	9	US-10-176-921-212	Sequence 212, App
32	83.5	4.4	747	9	US-10-137-865-212	Sequence 212, App
33	83.5	4.4	747	9	US-10-140-474-212	Sequence 212, App
34	83.5	4.4	747	9	US-10-142-431-212	Sequence 212, App
35	83.5	4.4	747	9	US-10-143-114-212	Sequence 212, App
36	83.5	4.4	747	9	US-10-140-002-212	Sequence 212, App
37	83.5	4.4	747	9	US-10-142-419-212	Sequence 212, App
38	83.5	4.4	747	9	US-10-123-263-212	Sequence 212, App
39	83.5	4.4	747	9	US-10-142-423-212	Sequence 212, App
40	83.5	4.4	747	9	US-10-121-050-212	Sequence 212, App
41	83.5	4.4	747	9	US-10-141-753-212	Sequence 212, App
42	83.5	4.4	747	9	US-10-143-032-212	Sequence 212, App
43	83.5	4.4	747	9	US-10-123-108-212	Sequence 212, App
44	83.5	4.4	747	9	US-10-123-236-212	Sequence 212, App
45	83.5	4.4	747	9	US-10-123-261-212	Sequence 212, App

ALIGNMENTS

RESULT 1

US-09-999-672-10

; Sequence 10, Application US/099999672

; Patent No. US20020127655A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GM1-SPECIFIC ALPHAL-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/09/999,672

; PRIOR FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: US/09/298,886

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; US-09-999-672-10

Query Match 100.0%; Score 1896; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.9e-188;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQQRIVKLPISKEPLMTTQSSGNTESPENRRDSEOHNGELRGMTINSIGLGNQM 60

|||||

Db 1 LQQRIVKLPISKEPLMTTQSSGNTESPENRRDSEOHNGELRGMTINSIGLGNQM 60

|||||

QY 61 GEYATLALARMNRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNVHLNDWKEER 120

|||||

Db 61 GEYATLALARMNRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPQNVHLNDWKEER 120

|||||

QY 121 YRHLPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSTFV 180

|||||

Db 121 YRHLPGHVFRTGYPCSWTFYHHLRPEILKEFTLHDHVREAAQAFRLGLRVNGSQSTFV 180

|||||

QY 181 GVHVRGDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240

|||||

Db 181 GVHVRGDYVHVMNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240

QY 241 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 300
Db 241 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 300
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353
Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 2
US-10-040-863-10
; Sequence 10, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-10

Query Match 100.0%; Score 1896; DB 12; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.9e-188; Indels 0; Gaps 0;
Matches 353; Conservative 0; Mismatches 0;
QY 1 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 1 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 60
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQSPSTFV 180
Db 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQSPSTFV 180
QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
QY 241 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 300
Db 241 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 300
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353
Db 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353

RESULT 3
US-09-999-672-8
; Sequence 8, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886

; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-999-672-8
Query Match 100.0%; Score 1896; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.4e-188; Indels 0; Gaps 0;
Matches 353; Conservative 0; Mismatches 0;
QY 1 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 28 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 83 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 147
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQSPSTFV 180
Db 148 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQSPSTFV 207
QY 181 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 240
Db 208 GVHVRGDYVHVMPNVKGVVADRGYLEKALDMFRARYSSPVFVVTSGMAWCRENINAS 267
QY 241 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 300
Db 268 RGVVFNAGIEGSPAKDFALLTQCNHTIMTIGTGGIWAAYLAGGDTIYLANVTLDPSPF 327
QY 301 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 353
Db 328 LKVKPEAAFLPEWVGIPADLSPLLKALTPACPRSHFLKAKGVTCYVAGRAF 380

RESULT 4
US-10-040-863-8
; Sequence 8, Application US/10040863
; Patent No. US20020137165A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/10/040,863
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/298,886
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-040-863-8
Query Match 100.0%; Score 1896; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.4e-188; Indels 0; Gaps 0;
Matches 353; Conservative 0; Mismatches 0;
QY 1 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 60
Db 28 LQORIVKLOPLSEKELPMTQSSGNTESPENRRDSEQHNGELRGMTINSIGRLGNQM 87
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 88 GEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKKIPWQNYHLNDWMEER 147
QY 121 YRHIPGHFVRETGYPCSWTFYHHLRPEILKEFTLHDHVREAAQFLRGLRVNGSQSPSTFV 180

Db 148 YRHIFGHEVFRFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPFV 207
QY 181 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINAS 240
Db 208 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINAS 267
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSPF 300
Db 268 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSPF 327
QY 301 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAGVTCYVAGRAF 353
Db 328 LKVFPEAAFLPEWVGIPADLSPLKALTPACPRSHFHLKAGVTCYVAGRAF 380

RESULT 5

US-09-051-034A-2

; Sequence 2, Application US/09051034A

; Patent No. US20010055584A1

; GENERAL INFORMATION:

; APPLICANT: MCKENZIE, IAN FAROUHAR CAMPBELL

; APPLICANT: SANDRIN, MAURO SERGIO

; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

; TITLE OF INVENTION: GLYCOSYLTRANSFERASE

; FILE REFERENCE: 30562.60SWO

; CURRENT APPLICATION NUMBER: US/09/051.034A

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: PCT/AU97/00492

; PRIOR FILING DATE: 1997-08-01

; PRIOR FILING DATE: 1996-08-21

; PRIOR APPLICATION NUMBER: P01402

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus Domesticus

US-09-051-034A-2

Query Match 75.4%; Score 1430; DB 10; Length 340;
Best Local Similarity 81.9%; Pred. No. 7.7e-140;
Matches 267; Conservative 26; Mismatches 19; Indels 14; Gaps 3;

QY 1 LQORIVKLOPLSEKELPMTTOMSSGNTESPMDRDSQHGNGELRGMTINSIGRLGNQM 60
Db 28 LQORIVKLOPLSEKELPMTTOMSSGNTESPMDRDSQHGNGELRGMTINSIGRLGNQM 73
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 74 GEYATLALARMNGREAFIPPEMHSTLAPIFRITLPVLHASTARRIPWQNYHLNDWMEER 133
QY 121 YRHIPGHEVFRFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPFV 180
Db 134 YRHIPGHEVFRFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPFV 193
QY 181 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINAS 240
Db 194 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINAS 253
QY 241 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSPF 300
Db 254 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSPF 313
QY 301 LKVFPEAAFLPEWVGIPADLSPLK 326
Db 314 LKLFKPEAAFLPEWVGIPADLSPLK 339

RESULT 6

US-09-999-672-11

Query Match 72.6%; Score 1376; DB 12; Length 344;
Best Local Similarity 78.9%; Pred. No. 3.1e-134;

; Sequence 11, Application US/09999672
; Patent No. US20020127655A1
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/999,672
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US/09/298,886
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-999-672-11

Query Match 72.6%; Score 1376; DB 10; Length 344;
Best Local Similarity 78.9%; Pred. No. 3.1e-134;
Matches 258; Conservative 26; Mismatches 31; Indels 12; Gaps 3;

QY 1 LQORIVKLOPLSEKELPMTTOMSSGNTESPMDRDSQHGNGELRGMTINSIGRLGNQM 60
Db 28 VQORLAKIQAM--WELPY-----QIPVLASTSKALGPSQLRGWMTINAIGRLGNQM 76
QY 61 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWQNYHLNDWMEER 120
Db 77 GEYATLALARMNGRLAFIPASMHNALAPIERISLPVLHSDTAKKIPWQNYHLNDWMEER 136
QY 121 YRHIPGHEVFRFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPF 179
Db 137 YRHIPGHEVFRFTGYPSCSWTFYHHLRPEILKEFTLHDHVREEAQAFRLGLRVNGSQSPF 196
QY 180 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINA 239
Db 197 GVHVRGDDYVHVMPNWKGVYADRGYLEKALDMFRARYSSPVFVTSNGMAWCARENINA 256
QY 240 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSP 299
Db 257 RGDVVFAGNGIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLAGGDTIYLANVTLPDSP 316
QY 300 FLKVFPEAAFLPEWVGIPADLSPLK 326
Db 317 FLKVFPEAAFLPEWVGIPADLSPLK 343

RESULT 7

US-10-040-863-11

; Sequence 11, Application US/10040863

; Patent No. US20020137165A1

; GENERAL INFORMATION:

; APPLICANT: Eric H. Holmes et al.

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE

; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES

; FILE REFERENCE: 8511-029

; CURRENT APPLICATION NUMBER: US/10/040,863

; PRIOR FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: 09/298,886

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 344

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-040-863-11

Wed May 28 09:21:53 2003

Db 51 PVAIFCLAGTVPHPNADSDCPKH-PASFSGTWITYPDGRFCNQMGQYATLLAQLNGRQ 109
 Qy 77 AFIPASHMNALAPRISLPLVHSDTAKKIPWQNYHLNDWMEERYRHPGHVFRFTGYPC 136
 Db 110 AFIOFAMHAYLAPVFTITPLVLAPEVDHAPWRELEHDMSEDYAHLKEPWLKLTGPPC 169
 Qy 137 SWTFYHHLRPEILKEFTLHDHVRREAQAFGLGRV--NGSOPSTFVGVHVRGDDVHVMP 194
 Db 170 SWTFPHHLRQIRSEFTLHDHVRREAQAFGLGRV--NGSOPSTFVGVHVRGDDVHVMP 229
 Qy 195 NVWKGWVADRGYLEKALDMFRARYSSPVVTSNGMWCNENINASRGDVVFAGNGIEGS 254
 Db 230 KWKGVGWDGRYLOQDMFRARYEAPVFTVTSNGMWCNENINASRGDVVFAGNGIEGS 289
 Qy 255 PAKDRALLTOCHNTTMTTGTGFWAAYLAGGDTIYLANFTLTPSSFLKIFKPEAAFLPEW 314
 Db 290 PAKDRALLTOCHNTTMTTGTGFWAAYLAGGDTIYLANFTLTPSSFLKIFKPEAAFLPEW 349
 Qy 315 VGIPADLSPL 324
 Db 350 VGINADLSPL 359
 RESULT 13
 US-09-848-838-2
 ; Sequence 2, Application US/09848836
 ; Patent No. US20020037570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Taylor, Diane
 ; APPLICANT: Wang, Ge
 ; APPLICANT: Palacic, Monica
 ; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
 ; FILE REFERENCE: 07254-061002
 ; CURRENT APPLICATION NUMBER: US/09/848,838
 ; PRIOR FILING DATE: 2001-05-03
 ; PRIOR APPLICATION NUMBER: US 09/433,598
 ; PRIOR FILING DATE: 1999-11-02
 ; PRIOR APPLICATION NUMBER: US 60/107,268
 ; PRIOR FILING DATE: 1998-11-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 300
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 US-09-848-838-2

Query Match 7.4%; Score 141; DB 10; Length 300;
 Best Local Similarity 24.0%; Pred. No. 2.9e-06;
 Matches 81; Conservative 33; Mismatches 125; Indels 98; Gaps 14;
 Qy 54 GRLGNMGCEYA-----TLFALARMNGRLAFIPASMHNALAPRIS-- 94
 Db 10 GGLGNMQFYAFKSLQKHLNTPVLLDTSDWSNKKWQLELPIDLPYANAKETAIAM 69
 Qy 95 --LPVLHSDTAKKIPWQNYHLNDWMEERYRHI-PGHVFRFTGYPCSWTFYHHLRPEILKE 151
 Db 70 QHLPKLVDRDALKYIGFDVRSQIVFPEYEPKLPKSLRTYFFGYFQDPYFDALSSLIKOT 129
 Qy 152 FTL-----HDHVRREAQAFGLGRVNGSOPSTFVGVHVRGDDVHVMPNVKMGVYA 202
 Db 130 FTLPPLPPPPNNKNNKEEYQKLS--LILAKNSVF--VHVRGDDVGI-----GCOL 179
 Qy 203 DRGYLEKALDMFRARYSSPVVTSNGMWCNENINASRGDVVFAGNGIEGSPAKDFA-- 260
 Db 180 GIDYQKALEYMAKRYPNMELFV-----FCE-----DLKFTQNLDLGYPFTDMITR 225
 Qy 261 -----LITQCNHTTMTTGTGFWAAYLAGGDTIYLANFTLTPD-----SPFLKVEK 305
 Db 226 DKREEAYWMLMCSCKHGIIANSTYSWAAYL-----MEN---PEKIIIGPKHWF 275
 Qy 306 PEAFLPEWVGIPADLSPLKALTPACPRSHHLKAK 342

Db 276 HENILCKREWVKI-----ESHFEVKSQ 296
 RESULT 14
 US-10-214-524-25
 ; Sequence 25, Application US/10214524
 ; Publication No. US20030073142A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Swey-Shen Alex
 ; APPLICANT: Yang, Yong-Min
 ; APPLICANT: Barankiewicz, Theresa J.
 ; APPLICANT: Chen, Zhong
 ; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF
 ; FILE REFERENCE: IGE-00101.P.1.1
 ; CURRENT APPLICATION NUMBER: US/10/214,524
 ; CURRENT FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 60/312,120
 ; PRIOR FILING DATE: 2001-08-13
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 25
 ; LENGTH: 496
 ; TYPE: PRT
 ; ORGANISM: Cat (Felis catus)
 US-10-214-524-25
 Query Match 5.2%; Score 99; DB 9; Length 496;
 Best Local Similarity 25.0%; Pred. No. 0.14;
 Matches 55; Conservative 21; Mismatches 84; Indels 60; Gaps 12;
 Qy 126 GHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVRREAQAFGLGRVNGSOPSTFVGVHVR 185
 Db 249 GEWVSQKTYTCQVTY-----QGFTEFHARKCTESDPRGSYILSPSPSL----- 293
 Qy 186 RGQYVHVMPNVKMGVADRGYLEKALDMFRARYSSPVVTSNGM--AWCRENINASRGD 243
 Db 294 -DLVYHSPKI-TCLVVD-----LANTDGMILTWSREN----- 324
 Qy 244 VVFAGNGIEGSPAKDFAALLTCNHTTMTTGTGFWAAYLAGGDTIYLANFTLTPDSP--F 300
 Db 325 ---GESVHPDP---MVKQYNGTITVTSTLPVDAIDWVEGET-YOCKVTHPDLPKDIV 376
 Qy 301 LKVPK-PEAAFLPE-WVGIPADLSPLK-ALTPACPRSHF 337
 Db 377 RSIAGAPGRREPPEVYVFLPPEGEPEKTKDKVTLTCLIQNF 416

RESULT 15
 US-09-479-614-14
 ; Sequence 14, Application US/09479614
 ; Publication No. US20030013183A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine
 ; APPLICANT: Weber, Eric
 ; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
 ; FILE REFERENCE: P-1047
 ; CURRENT APPLICATION NUMBER: US/09/479,614
 ; CURRENT FILING DATE: 2000-01-07
 ; EARLIER APPLICATION NUMBER: 60/115,033
 ; EARLIER FILING DATE: 1999-01-07
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Felis catus
 US-09-479-614-14
 Query Match 5.1%; Score 97.5; DB 9; Length 431;
 Best Local Similarity 25.0%; Pred. No. 0.16;
 Matches 52; Conservative 19; Mismatches 78; Indels 59; Gaps 11;
 Qy 126 GHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVRREAQAFGLGRVNGSOPSTFVGVHVR 185

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.7916 Seconds

(without alignments)

2652.940 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQQRIVKLPSEKELPMTT.....RSHFHLKAKGVICYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

.Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1386.5	73.1	343	2 A56098	alpha(1,2)fucosylit
2	1291	68.1	354	2 B56392	beta-galactoside a
3	1080.5	57.0	373	2 A56392	beta-galactoside a
4	1051.5	55.5	365	2 A36047	galactoside 2-alpha
5	846	44.6	159	2 S46494	galactoside 2-alpha
6	581	30.6	142	2 S51582	galactoside 2-alpha
7	214.5	11.3	281	2 T44328	hypothetical prote
8	183	9.7	443	2 H87911	protein B0205.4 [i
9	170.5	9.0	363	2 T20745	hypothetical prote
10	154	8.1	348	2 T31916	hypothetical prote
11	153.5	8.1	335	2 T32294	hypothetical prote
12	139.5	7.4	383	2 T20572	hypothetical prote
13	138.5	7.3	500	2 T23068	hypothetical prote
14	133.5	7.0	299	2 H71976	hypothetical prote
15	131	6.9	365	2 T25309	probable alpha(1,2
16	127	6.7	388	2 T32307	hypothetical prote
17	127	6.7	392	2 T25334	hypothetical prote
18	124	6.5	353	2 T25390	hypothetical prote
19	120	6.3	625	2 T21051	hypothetical prote
20	119	6.3	434	2 T26275	hypothetical prote
21	118.5	6.2	381	2 T15140	hypothetical prote
22	117.5	6.2	365	2 T32253	hypothetical prote
23	111	5.9	395	2 T32309	hypothetical prote
24	109.5	5.8	371	2 T32592	hypothetical prote
25	109	5.7	277	2 T25307	hypothetical prote
26	98.5	5.2	317	2 T33887	hypothetical prote
27	96.5	5.1	355	2 T34405	hypothetical prote
28	93.5	4.9	348	2 T02798	hypothetical prote
29	93	4.9	597	2 G87258	acyl-CoA dehydroge

ALIGNMENTS

RESULT 1

A56098

alpha(1,2)fucosyltransferase Sec2, long form - human

C:Species: Homo sapiens (man)

C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 21-Jul-2000

C:Accession: A56098

R:Keilly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.

J. Biol. Chem. 270, 4640-4649, 1995

A:Title: Sequence and expression of a candidate for the human Secretor blood group al

tes with the non-Secretor phenotype.

A:Reference number: A56098; MUID:95181460; PMID:7876235

A:Accession: A56098

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-343 <REL>

A:Cross-references: GB:U17894; NID:9687618; PIDN:AAC24453.1; PID:9687619

C:Genetics:

A:Gene: GDB:FUT2; SE

A:Cross-references: GDB:120619; OMIM:182100

A:Map position: 19q13.3-19q13.3

C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein

Query Match

Best Local Similarity 79.1%; Score 1386.5; DB 2; Length 343;

Matches 258; Conservative 26; Mismatches 31; Indels 11; Gaps 2;

OY 1 LQQRIVKLPSEKELPMTTOMSSGNTSEPMRDSQHGNGELRGMTINSTIGRLGNQM 60

Db 28 VOORLAKIQAM--WELPV-----QIFVLASTSKALGPSOLRGMTINAIIGRLGNQM 76

OY 61 GEVATLFLARMNGRLAFIPASHMNAIATFRISLPVLSHSDTAKKIPWQNYHLNDWKEER 120

Db 77 GEAYILYALAKWNGRPAFPAQMHSTLAFITPLVLSHSDTAKKIPWQNYHLNDWKEE 136

OY 121 YRHIPGHFVRFTGYPCSWTFYHHLRPEILKEFTLHDHVREREAQAFRLGRVNGSQPSTFV 180

Db 137 YRHIPGEYVRFYGYPCSWTFYHHLRQEILOEFTLHDHVREREAQAFRLGRVNGSRPGT 196

OY 181 GVHVARGDYVHVMPNWKGVADRGYLEKALDMFRARYSSPVFVTSNGMWCRENINAS 240

Db 197 GVHVRRGDYVHVMPNWKGVADRRYLOQALDWFARYSSLIFFVTSNGMWCRENIDS 256

OY 241 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTTGTGGINAYLAGDGTIYLANITLPDSPF 300

Db 257 RGDVVFAGNGIEGSPAKDFALLTQCNIHMTTGTGGINAYLAGDGTIYLANITLPDSPF 316

OY 301 LKVEKPEAAFLPEWVGIPADLSPLLK 326

Db 317 LKIFKPEAAFLPEWVGIAADLSPLLK 342

RESULT 2

```
B56392
beta-galactoside alphas,2-fucosyltransferase II - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: B56392
R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: B56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-354 <HIT>
A:Cross-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357
C:Keywords: transmembrane protein

Query Match 58.1%; Score 1291; DB 2; Length 354;
Best Local Similarity 83.9%; Pred. No. 1.4e-100; Indels 0; Gaps 0;
Matches 235; Conservative 20; Mismatches 25;

QY 46 GMEFTNSIGRLGNOMGEYATLALARMNGRLAFIPASMHNALAPIFRISLPVLHSDTAKK 105
DB 71 GWTINAMGLGNOMGEYATLALAKENGRLPAYIPAQMHSTLAPIFRISLPVLHSDTAKK 130
QY 105 IPQNYHLNDWMEERYRHIPGHVFRTPGCSWTFYVHHLRPEILKEFTLHDHVRERAAQAF 165
DB 131 VPQNYHLNDWMEERYRHIPGVFVRLTGYPCSWTFYVHHLRPEILKEFTLHDHVRERAAQAF 190
QY 166 LRLGLRVNGSQPSTFVGVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRARYSSPFFVY 225
DB 191 LRLGLRVNGSRPSTFVGVHVRGDDYVHVMNVKGVVADRGYLEKALDMFRAPTPPFFVY 250
QY 226 TSNMGAWCRENINASRGDVVAGNGIEGSPAKDFALLTCNHTIMTIGFGIWAAYLAGG 285
DB 251 TSNMGAWCRENIDASRGDVVAGNGIEGSPAKDFALLTCNHTIMTIGFGIWAAYLAGG 310
QY 286 DTIYLANVTLPDSPFTLVKFKPEAAFLPEWVGIPADLSPL 325
DB 311 DTIYLANVTLPDSPFTLVKFKPEAAFLPEWVGITANMGAL 350

RESULT 3
A56392
beta-galactoside alphas,2-fucosyltransferase I - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: A56392
R:Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphas
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: A56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <HIT>
A:Cross-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355
C:Keywords: transmembrane protein

Query Match 57.0%; Score 1080.5; DB 2; Length 373;
Best Local Similarity 63.3%; Pred. No. 6.9e-83;
Matches 200; Conservative 44; Mismatches 51; Indels 21; Gaps 3;

QY 16 LPMFTTOMS-----SQNTESPENMRDSEOHNGELRGMTINSIGRLGNOMGEYATLALA 70
DB 58 LPMFTTOMS-----SQNTESPENMRDSEOHNGELRGMTINSIGRLGNOMGEYATLALA 103
QY 71 RNRGLAFTPASMHNALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHFVR 130
DB 104 QLNREAFILPAMHAALAPVFRITLPLAPEVNRRTSKQLLDHOMSEYSLDPPFLK 163
QY 131 FTGYPGCSWTFYVHHLRPEILKEFTLHDHVRERAAQAFRLGLRV--NGSQPSTFVGVHVRGDD 188
DB 164 FTGYPGCSWTFYVHHLRPEILKEFTLHDHVRERAAQAFRLGLRGTRTPFTFVGVHVRGDD 223

QY 189 YVHVMPNWKGVVADRGYLEKALDMFRARYSSPFFVVTNSMGAWCRENINASRGDVVYFAG 248
DB 224 YLQVMPQPNKGVVGDRAVYLAQDMFRARHEAPFVVTNSMGAWCRENIDASRGDVWFAG 283
QY 249 NGIEGSPAKDFALLTCNHTIMTIGFGIWAAYLAGGDTIYLANVTLPDSPFTLVKFKPEA 308
DB 284 NGLSSPAKDFALLTCNHTIMTIGFGIWAAYLAGGDTIYLANVTLPDSPFTLVKFKPEA 343
QY 309 AFLPEWVGIPADLSPL 324
DB 344 AFLPEWVGINADLSPL 359

RESULT 4
A36047
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - human
C:Species: Homo sapiens (man)
C>Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 10-Sep-1997
C:Accession: A36047
R:Larsen, R.D.; Ernst, L.K.; Nair, R.P.; Lowe, J.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 6674-6678, 1990
A:Title: Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-g
A:Reference number: A36047; MUID:90370848; PMID:2118655
A:Accession: A36047
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-365 <LAR>
A:Cross-references: GB:M35531; NID:g183887; PID:g306830
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 55.5%; Score 1051.5; DB 2; Length 365;
Best Local Similarity 63.3%; Pred. No. 1.8e-80;
Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;

QY 25 GNTESPENMRDSEOHNGELRGMTINSIGRLGNOMGEYATLALARMNGRLAFIPASMH 84
DB 59 GTAMPNASSCCPQH-PASLSGTWTVYVNGRFGNMGQYATLLAQLNGRRAFILPAMH 117
QY 85 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYRHIPGHVFRTPGCSWTFYVHHL 144
DB 118 AALAPVFRITLPLAPEVDSRTPNRELQLHDWKEEYADLRDLPELKLSGFPSCWTFPHHL 177
QY 145 RPEILKEFTLHDHVRERAAQAFRLGLRV--NGSQPSTFVGVHVRGDDYVHVMNPVWKGVVA 202
DB 178 REQIRERFTLHDHVRERAAQAFRLGLRGTRTPNRELQLHDWKEEYADLRDLPELKLSGFPSCWTFPHHL 237
QY 203 DRGYLEKALDMFRARYSSPFFVVTNSMGAWCRENINASRGDVVAGNGIEGSPAKDFALL 262
DB 238 DSAYLRQAMDWRARHEAPVVTNSMGAWCRENIDASRGDVVAGNGIEGSPAKDFALL 297
QY 263 TQCNHTIMTIGFGIWAAYLAGGDTIYLANVTLPDSPFTLVKFKPEAAFLPEWVGIPADLS 322
DB 238 TQCNHTIMTIGFGIWAAYLAGGDTIYLANVTLPDSPFTLVKFKPEAAFLPEWVGIPADLS 357
QY 323 PLLKALTP 330
DB 358 PLWTLAKP 365

RESULT 5
S46494
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) B - rat (fragment)
N:Alternate names: alpha-1,2-fucosyltransferase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Dec-1994 #sequence_revision 23-Feb-1996 #text_change 08-Oct-1999
C:Accession: S46494
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46494
A:Molecule type: mRNA
```

A:Residues: 1-159 <PIA>
A:Cross-references: EMBL:L26010; NID:g414816; PIDN:AAB41515.1; PID:g414817
C:Genetics: FTB
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 44.6%; Score 846; DB 2; Length 159;
Best Local Similarity 99.4%; Pred. No. 1.1e-63;
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 195 NWKGVADRGYLEKALDMFRARYSSPVVTSNGMAWCNINASRGDVVFNAGNGIEGS 254
DB 1 NWKGVADRGYLEKALDMFRARYSSPVVTSNGMAWCNINASRGDVVFNAGNGIEGS 60
QY 255 PAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLYNTLPDSPFLKVKPEAAFLPEW 314
DB 61 PAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLYNTLPDSPFLKVKPEAAFLPEW 120
QY 315 VGIPADLSPILKALTPACPRSHFHLKAKGVTCYVAGRAF 353
DB 121 VGIPADLSPILKALTPACPRSHFHLKAKGVTCYVAGRAF 159

RESULT 6
S51582
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1995 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C:Accession: S51582; S46493
R:Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
submitted to the EMBL Data Library, November 1993
A:Reference number: S51582
A:Accession: S51582
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-142 <PIA>
A:Cross-references: EMBL:L26009; NID:g414814; PIDN:AAB41514.1; PID:g554438
R:Plau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially expressed in the rat
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46493
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 9-135 <PI2>
A:Cross-references: EMBL:L26009
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 30.6%; Score 581; DB 2; Length 142;
Best Local Similarity 73.2%; Pred. No. 1.5e-41;
Matches 104; Conservative 15; Mismatches 21; Indels 2; Gaps 1;

QY 135 PCSWTFYHHLRPEILKFTLHDHVREAAFLRGLRYN--GSQSTFVGHVRRGDYVHV 192
DB 1 PCSWTFYHHLRPEILKFTLHDHVREAAFLRGLRYN--GSQSTFVGHVRRGDYVHV 60
QY 193 MPNVKGVADRGYLEKALDMFRARYSSPVVTSNGMAWCNINASRGDVVFNAGNGIE 252
DB 61 MPNVKGVADRGYLEKALDMFRARYSSPVVTSNGMAWCNINASRGDVVFNAGNGIE 120
QY 253 GSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLYNTLPDSPFLKVKPEAAFLPEW 314
DB 121 GSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLYNTLPDSPFLKVKPEAAFLPEW 159

RESULT 7
T44328
hypothetical protein wbla [imported] - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
C:Accession: T44328
R:Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda, Y.
Gene 237, 321-332, 1999

A:Title: The genes responsible for O-antigen synthesis of Vibrio cholerae O139 are cl
A:Reference number: T22749; MUID:99453293; PMID:10521656
A:Accession: T44328
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-281 <YAM>
A:Cross-references: EMBL:AB012957; NID:g4115688; PIDN:BAA33632.1; PID:g3721682
A:Experimental source: strain O22
C:Genetics:
A:Note: wbla
C:Superfamily: Vibrio cholerae hypothetical protein wbla

Query Match 11.3%; Score 214.5; DB 2; Length 281;
Best Local Similarity 26.5%; Pred. No. 2e-10;
Matches 79; Conservative 50; Mismatches 102; Indels 67; Gaps 14;

QY 47 MFTINSIGRLNQMGAYATLPALARMNG--RL---AFIPASMHNALAPIRISLPVLHS 100
DB 1 MIVMKISGGLNQLFQYAVGRAIAIQGVPLKLDVSAYKNYKLHNG---YRLDOFNINA 56
QY 101 DTAKKIPWQNYHL-----NDWMEERYRHI-----PGHFVRFT 132
DB 57 DIANED--EIFHLKSSNRLSRLRLRLGLWLNKNTYYAEKQRTIYDVSVFMQAPRY---LD 111
QY 133 GYPCSWTFYHHLRPEILKFTLHDHVREAAFLRGLRVNGSQSTFVGHVRRGDYVHV 192
DB 112 GYQWNEQYFSQIRAVLLOELWPNQPLSINAQA--HQIKI--QQTHAVSIHVRGDIYN- 165
QY 193 MPNVKGVADRGYLEKALDMFRARYSSPVVTSNGMAWCNINASRGDVVFNAGNGIE 252
DB 166 HPEIG--VLDIDYKRAVDYIKIEAPVFFVSNVAVWCKDNFNFDSPVFIETDTOTE 222
QY 253 GSPAKDFALLTQCNTHTMTIGTIGIWAAYLAGGDTIYLYNTLPDSPFLKVKPEAAFLPEW 314
DB 223 ---IDDLMLQCQOHNIIVANSFSSFWAAFLNSVDKIVIAPKTWAENPKGVKWPDS 277

RESULT 8
H87911
protein B0205.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: H87911
R:anonymous, the C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: H87911
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: GB:chr.I; PIDN:AA16988.1; PID:g3150470; GSPDB:GN00019; CESP:B020
A:Note: contains weak similarity to fucosyltransferases
C:Genetics:
A:Gene: B0205.4
A:Map position: 1

Query Match 9.7%; Score 183; DB 2; Length 443;
Best Local Similarity 20.6%; Pred. No. 1.6e-07;
Matches 72; Conservative 53; Mismatches 128; Indels 96; Gaps 11;

QY 56 LGNQMGAYATLPALARMNGRLAFIPA--SMHNA--LAPIRISLPVLHSDTAKKI----- 106
DB 103 IGNLMFOVAGLLSIARETGSILLIPSTTLRRAFDFTFNDSIQFVGEDLSRLAEDLN 162
QY 107 -----PWQYHNLNDWMEERYRHPGHVFTGYPCSWTFYHHLRPEILKE-FTLH 155
DB 163 ASKITLTSACCAVRLNLTILFNDSRI-----IERIDGYFQNFRRHPDSQKIVKLTFTM 216
QY 156 DHVPEEAQAF-----LRLGLR 170

Db 217 DPVRKRVSEWNIYWIHPTNHRKKKEKSTVSIFSEVTFQLKRVDFLENVGSIETVRNAR 276
Qy 171 V-----NGSQ-----PSTFVGVHVRGDYVHV--MPNVWKGWVADRGYLEKAL 211
Db 277 VIETNVANDQALELPEEDAFKATNMVGVHIRHGMDSISNRSNRHGHVDVPTPIEYIKRAI 336
Qy 212 DMFRARYSPFVVT'SNGMAWRENINASRGDVFVAGNGIEGSPAKDFALLTQCNTHTMT 271
Db 337 QOISKIYENAFIICSDNVAVARRNLKLGKETLHEF---CPGPREVDMAILKSCDSVIIS 393
Qy 272 IGTGCIWAAYL---AGGDTIYLANVTLDPSPFLKVPKPEAAFLPEWGI 317
Db 394 TCTGFWWSAYLVNWNASPDVYKYKHPAPGSMVEKMTNKTETPLKSWTAL 442

RESULT 9
T20745
hypothetical protein F11A5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20745
R:Gardner, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19319
A:Accession: T20745
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-363 <WIL>
A:Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5
A:Experimental source: clone F11A5
C:Genetics:
A:Gene: CESP:F11A5.5
A:Map position: 5
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 9.0%; Score 170.5; DB 2; Length 363;
Best Local Similarity 23.4%; Pred. No. 1.4e-06;
Matches 73; Conservative 45; Mismatches 101; Indels 93; Gaps 17;

Qy 52 SIGRLNGMGEYATLAFALARMGRLA--RIPASMH-----NALAPIFRISLPVLHSDT 102
Db 78 SSSRLGNHLFELASVLSISRELQVPTFTIENCYHEKMWEDSNLIPGLMNHFLINGSV 137
Qy 103 AKKIPQWNI-----LNDMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEF 152
Db 138 PSVGRVKVHKQKCTFDDPSLLDNYDEYHLTG-----THYQ-SWKYFSHMRNELI--- 188
Qy 153 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVPNVWKGV-VADRGYLEKA 210
Db 189 ---GYLKTENTY---MDLPKSGENTFITCVHVRRGDFLRV-----GFHVADENFIRSS 236
Qy 211 LDMFRARYSPFVVT'SNGMAWRENINASRGDVFVAGNGIE----- 252
Db 237 LNLISQVAKRANTAT-----VFGDDYEFMDSLNRKTSINAFVSQ 278
Qy 253 GSPAKDFALL-TQCNHTMTI--GTGCIWAAYLAGGDTIYLANVTLDPSPFLK----- 302
Db 279 NSPADLLYAKSCDVVLIITAASHSTFGWGMGYPSKGNRY--YT---DIQTKDWILETG 333
Qy 303 VFKEPAFLPEW 314
Db 334 EFSIEDYILPHW 345

RESULT 10
T31916
hypothetical protein C17A2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T31916
R:Sammons, L.; Wohldmann, P.
submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid C17A2.
A:Reference number: Z21098
A:Accession: T31916
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-348 <SAM>
A:Cross-references: EMBL:AF016654; PIDN:AAB66030.1; GSPDB:GN00020; CESP:C17A2.4
A:Experimental source: strain Bristol N2; clone C17A2
C:Genetics:
A:Gene: CESP:C17A2.4
A:Map position: 2
A:Introns: 94/3; 133/3; 168/2; 272/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.1%; Score 154; DB 2; Length 348;
Best Local Similarity 22.8%; Pred. No. 3.2e-05;
Matches 63; Conservative 53; Mismatches 102; Indels 58; Gaps 15;

Qy 52 SIGRLNGMGEYATLAFALARMGR--LAFIPASMHNALAPIFRISLP-----VLHSDT 102
Db 36 SPSRLGNHIFEFASLFLSERLHRTPLFLVNEFHOKMLDETAKVMFGLVEKFTVINGSL 95
Qy 103 AKKI---PHQ-----NYHLNDMEERYRHIPGHEVFTGYPCSWTFYHHLRPEILKEF 152
Db 96 PKSIKOTPFQKVCCHRENPEILEKIDKYLHLTGMYQ-----SWKYPNMRQQL--- 146
Qy 153 TLHDHVREEAQAFRLGLRVNGSQPSTFVG-VHVRGDYVHVPNVWKGV-VADRGYLEKAL 211
Db 147 ---DFLDDSSQDF--GNLPSRNOHTVTCVHARRGFDVY-----GFOAADPDFIRNSV 195
Qy 212 DMFRARYSPV-----FVVT'SNGMAWCR---EN--NASRGDVFVAGNGI--EGSPA 256
Db 196 KYTAENFIPEIEYKVKVRKVIFGDDLEFMRSLFENSFVSTDEPEYMFPAEYISQNSPA 255
Qy 257 KOFALLTQ--CNHTMTI--GTGCIWAAYLAGGDTIY 289
Db 256 EDLTYSKQNCDIVLISAPKSTFGWGMGYPSKGNKV 291

RESULT 11
T32294
hypothetical protein K06H6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T32294
R:David, M.; Wohldmann, P.; Bauer, C.; Clarke, K.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid K06H6.
A:Reference number: Z21147
A:Accession: T32294
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-335 <DAV>
A:Cross-references: EMBL:AF024500; PIDN:AAB70365.1; GSPDB:GN00023; CESP:K06H6.6
A:Experimental source: strain Bristol N2; clone K06H6
C:Genetics:
A:Gene: CESP:K06H6.6
A:Map position: 5
A:Introns: 17/1; 32/2; 58/1; 111/3; 150/3; 184/2; 218/1; 243/3; 272/1; 303/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.1%; Score 153.5; DB 2; Length 335;
Best Local Similarity 22.8%; Pred. No. 3.3e-05;
Matches 77; Conservative 50; Mismatches 118; Indels 93; Gaps 19;

Qy 24 SGNTESEPMRDRSDBOHNGELRGMTINSIGRLNGMGEYATLAFALARMGRLA--PIPA 81
Db 27 SATAEIPLRQESFLIMKKRLSSRMA--PIARLGNHMFELAAVLGIGSRMLNRTATFIED 84
Qy 82 SMHNALAPIFRISLP-----VLHSDTAKKIPW--QNYHLN-----DWMER 120
Db 85 EIYRRMTESSEKAIPLGLVGQFEILNG---KVPLYIKTKLNTKCCVFDPLTHEHNDE 140

hypothetical protein F41D3.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22068
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19508
A:Accession: T22068
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-500 <WIL>
A:Cross-references: EMBL:Z81537; PIDN:CAB04377.1; GSPDB:GN00019; CESP:F41D3.6
A:Experimental source: clone F41D3
C:Genetics:
A:Gene: CESP:F41D3.6
A:Map position: 1
A:Introns: 23/1; 59/1; 112/3; 150/3; 222/1; 278/1; 308/3; 326/3; 421/3

Query Match 7.3%; Score 138.5; DB 2; Length 500;
Best Local Similarity 23.4%; Pred. No. 0.001;
Matches 74; Conservative 48; Mismatches 105; Indels 89; Gaps 18;

Qy 47 MTTIVYGGKLNQFVLISLLGIARKLKRIA----VENSDDPVLQSLSFLNOKLPRI 103
| | | | | : | : | | | : | : | : | : | : | :
Db 49 MTTIVYGGKLNQFVLISLLGIARKLKRIA----VENSDDPVLQSLSFLNOKLPRI 103

Qy 93 ---ISLPLVHSDTAK-----KIPWNVHLNDMVEERYRHPIGHVFRVTGTPCSWTFF 141
| | | : | : | : | : | : | : | : | : | : | : | : | :
Db 104 EQVISVETEPSTRTFAISSDCRYELSDNLAD--ESKFLVIEGHVFQ-----SYKYF 155

Qy 142 HHLRPETILKEFTLHDHVREAAQAFRLGLRVNGSQPSFTFGVHVRGDYV---HVMNVW 197
: | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 156 ADMKLSI-KEWL---KPDEPEKPRMKISTESORHK-TCVHVRGDFLDEQH----- 203

Qy 198 KGVVADRGLYEKALDMFRARYSSPVVVTSNGMAWCRENINASRGDVYVFNAGLEGSPAK 257
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 204 --AGTDSNYTISAIDHLSLESLHGVIIFIM-SNDPKWKVHI-AOHLDY-----QK 248

Qy 258 DFALL-----TOCNHTIMTI--GTFGIWAAYLAGGDT-IYLANY-TLPDS 298
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 249 DIRIKTMEDAIDDLHFSSQICYDSVLITAPSSTFGWIGYLSKNQSAVYRDIRETKDQ 308

Qy 299 PFLVKFKPERAFLEPW 314
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 309 VOLQMTK-EDFYPTW 323

RESULT 14
H71976
Probable alpha(1,2)fucosyltransferase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000
C:Accession: H71976
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.
Nature 397, 176-180, 1999
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71976
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <ARN>
A:Cross-references: GB:AEO01447; GB:AEO01439; NUD:g4154503; PIDN:AAD05659.1; PID:g415
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0086
C:Superfamily: Vibrio cholerae hypothetical protein wbla

Query Match 7.0%; Score 133.5; DB 2; Length 299;
Best Local Similarity 22.9%; Pred. No. 0.0014;
Matches 76; Conservative 38; Mismatches 129; Indels 89; Gaps 13;

```

QY 54 GRIGNQGEYA-----TLFALARNGRLAIFASMHNALAPIFRIS-- 94
Db 10 GGLGNMQFOYAFAKSLQKHSNTPLVLDITSFQGSNKMQLGPFIDLPYASAKEIATAM 69
QY 95 --LPVLHSDTAKKIPQNVYHLNDWMEEYRHI--PGHFVRFCTGYPCSWTFYHHLRPEILKE 151
Db 70 QHLPKLVYDALKYNGFDRYSQEIFVEYEPKLLKPSRLTYFYGFQDPYDFDAISSLIQQT 129
QY 152 FTL-----HDHVREEAQAFLRGLR-VNGSQPSTFVGHVHVRRGDIYHVMPNVWKGVVADRQ 205
Db 130 FTLPPLPPPPENGNNKKKEEYHRKLSLLAAKNSVF--AHIRRGDYVGI-----GCQLGID 181
QY 206 YLEKALDMPFARYSPFVVVTSNGMAWCORENINASRGDVFVAGNLEGSAPKDEFA----- 260
Db 182 YQRKAVEYMAKRVPMELFV-----FCZ-----DLKFTONLDLGFPFMDMTTRDKD 227
QY 261 -----LLTCNHNHIMTIGTFGIWAAYLAG--GDTIYLANITYLPDPSFLKVFVKPEAAF 310
Db 228 EEAYDMLLMQSCKHGIANSIYSWAAVLINNPGRKII-----GPKHWFGLGCHENIL 279
QY 311 LPFWGIPDALPSLLKALTPACPRSHFLKAK 342
Db 280 CREWKVI-----ESHFEVKSQ 295

RESULT 15
T25309
hypothetical protein T26E4.4 - Caenorhabditis elegans
C:Species:Caenorhabditis elegans
C:Date:15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T25309
R:McMurray, A.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z20013
A:Accession: T25309
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 SWIL>
A:Cross-references: EMBL:Z61132; PIDN:CAB03434.1; GSPDB:GN00023; CESP:T26E4.4
A:Experimental source: clone T26E4
C:Genetics:
A:Gene: Cesp:T26E4.4
A:Map position: 5
A:Introns: 50/1; 80/2; 140/3; 179/3; 218/2; 253/1; 278/1; 306/3; 325/1; 342/3
C:Superfamily:Caenorhabditis elegans hypothetical protein C14C6.3

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.3317 Seconds
(without alignments)
1417.103 Million cell updates/sec

Title: US-10-040-863-10

Perfect score: 1896

Sequence: 1 LQRIYKLPFLSEKELPMTT.....RSHFHLKAKGVTCYVAGRAF 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1430	75.4	340	1	FUT2_PIG
2	1293	73.5	344	1	FUT2_BOVIN
3	1386.5	73.1	343	1	FUT2_HUMAN
4	1338	70.6	368	1	FUT2_MOUSE
5	1291	68.1	354	1	FUT2_RABIT
6	1105.5	58.3	376	1	FUT1_MOUSE
7	1101.5	58.1	376	1	FUT1_RAT
8	1080.5	57.0	373	1	FUT1_RABIT
9	1051.5	55.5	365	1	FUT1_HUMAN
10	1050.5	55.4	365	1	FUT1_PIG
11	846	44.6	159	1	FUT2_RAT
12	168	8.9	365	1	YKQ7_CAEEL
13	91	4.8	442	1	VGLD_HSVBK
14	89	4.7	353	1	FNAB_PORGI
15	89	4.7	773	1	CDH_PBACH
16	89	4.7	2058	1	MT10_HUMAN
17	88.5	4.7	359	1	PTN7_RAT
18	88.5	4.7	402	1	VGLD_HSVBA
19	88.5	4.7	452	1	VGLD_HSVBA
20	88	4.6	217	1	GRB2_MOUSE
21	87	4.6	217	1	GRB2_HUMAN
22	87	4.6	438	1	NOO1_THETH
23	87	4.6	533	1	OUTD_EMENI
24	86.5	4.6	1112	1	CN3B_HUMAN
25	85.5	4.5	738	1	PL03_HUMAN
26	85.5	4.5	1928	1	LPH_RAT
27	84.5	4.5	344	1	RIBB_HELPY
28	84.5	4.5	383	1	AAT_THEAT
29	84.5	4.5	718	1	CTPC_MYCTU
30	84	4.4	1780	1	POIG_MVEV
31	83.5	4.4	273	1	TC1A_CAEEL
32	83.5	4.4	350	1	FMB2_FORGI
33	83.5	4.4	440	1	CYE_PARDE

RESULT 1
FUT2_PIG

ID FUT2_PIG STANDARD; PRT; 340 AA.

AC Q10982; Q29044; O19100;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-

DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)F 2)

DE Fucosyltransferase 2).

GN FUT2.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97468270; PubMed=9321466;

RA Meijerink E., Fries R., Voegeli P., Masabanda J., Wigger G.,

Stricker C., Neuenschwander S., Bertschinger H.U., Stranzinger G.,

Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6c11

are closely linked to the blood group inhibitor (S) and Escherichia

coli F18 receptor (ECF18R) loci.;

Mamm. Genome 8:736-741(1997).

RL [2]

RP SEQUENCE FROM N.A.

RA Conney S., Mountouris E., McKenzie I.F.C., Sandrin M.S.;

"Molecular cloning and characterization of the pig secretor type

alpha(1,2)fucosyltransferase.;"

Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

RL [3]

RP SEQUENCE OF 69-334 FROM N.A.

RA Petit J.M.;

Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE OF 69-80; 119-133 AND 316-334.

TISSUE=Submaxillary gland;

MEDLINE=96064742; PubMed=7592879;

Thurin J., Blaszyk-Thurin M.;

"Porcine submaxillary gland GDP-L-fucose: beta-D-galactoside

alpha-2-L-fucosyltransferase is likely a counterpart of the human

Secretor gene-encoded blood group transferase.;"

J. Biol. Chem. 270:26577-26580(1995).

RL -1- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +

alpha-L-fucosyl-1,2-beta-D-galactosyl-R.

CC -1- PATHWAY: Glycosylation.

CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND

FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC

CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SEC1) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: Y09882; CA71008.1; -.
DR EMBL: AF113532; AAD25351.1; -.
DR MGD: MGI:109374; Fut2.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 42 368 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 368 AA; 41464 MW; 4093E853EB37303B CRC64;

Query Match 70.6%; Score 1338; DB 1; Length 368;
Best Local Similarity 88.0%; Pred. No. 9.9e-106;
Matches 243; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 46 GMFTINSTGRGNQMGVATLALARMNGRLAFIPASHMNLAPIFRISLPVLHSDTAKK 105
DB 69 GMFTIRVRGRGNQMGVATLALARMNGRLAFIPASHMNLAPIFRISLPVLHSDTAKR 128
QY 106 IPWQNYHLNDWMEERYRHPGHEVFTGPGSWTGYHLRPEILKEFTLHDHVRERAAQAF 165
DB 129 IPWQNYHLNDWMEERYRHPGHEVFTGPGSWTGYHLRPEILKEFTLHDHVRERAAQAF 188
QY 166 LRGLRVNGSQSPSTFGVGVHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFFV 225
DB 189 LRGLRVNGSQSPSTFGVGVHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFFV 248
QY 226 TSNMGAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNTHTMTIGTFTGIIWAAYLAGG 285
DB 249 TSDDMWCRCSTASRGDVFAGNIEGSPAKDFALLTQCNTHTMTIGTFTGIIWAAYLAGG 308
QY 286 DTIYLANVTLPDSPPLKFKPEAAFLPEWVGIPADL 321
DB 309 DTIYLANVTQNSPHTVFKPEAAFLPEWVGIAADL 344

RESULT 5
FUT2_RABIT STANDARD; PRT; 354 AA.
AC Q10983;
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood
DE group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-
DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)
DE (Fucosyltransferase 2).
GN FUT2 OR SEC1 OR RFT-II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC (1,2)GALBETA- CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80225; CAA56512.1; -.
DR InterPro: IPR002516; GT_11.
DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 44 354 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6CCB4 CRC64;
Query Match 68.1%; Score 1291; DB 1; Length 354;
Best Local Similarity 83.9%; Pred. No. 8.9e-102;
Matches 235; Conservative 20; Mismatches 25; Indels 0; Gaps 0;
QY 46 GMFTINSTGRGNQMGVATLALARMNGRLAFIPASHMNLAPIFRISLPVLHSDTAKK 105
DB 71 GMFTINAMGRGNQMGVATLALAKENGPRAYIPAOHMSIAPIFRISLPVLHSDTAKR 130
QY 106 IPWQNYHLNDWMEERYRHPGHEVFTGPGSWTGYHLRPEILKEFTLHDHVRERAAQAF 165
DB 131 IPWQNYHLNDWMEERYRHPGHEVFTGPGSWTGYHLRPEILKEFTLHDHVRERAAQAF 190
QY 166 LRGLRVNGSQSPSTFGVGVHVRGDYVHVMPNWKGVVADRGYLEKALDMFRARYSSPFFV 225
DB 191 LRGLRVNGSRPSTFGVGVHVRGDYVHVMPNWKGVVADRGYLEQALDWFRAPTAPFFV 250
QY 226 TSNMGAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNTHTMTIGTFTGIIWAAYLAGG 285
DB 251 TSNMGAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNTHTMTIGTFTGIIWAAYLAGG 310
QY 286 DTIYLANVTLPDSPPLKFKPEAAFLPEWVGIPADL 325
DB 311 DTIYLANVTAPDSPPLKFKPEAAFLPEWVGITANMGAL 350
RESULT 6

QY	189	YVHVYPNWKGVADRGYLKALDMPFARYSSVFVVTSNGMAWCARENINASRGDVVFAG	246
Dd	225	YLEYPPNRWKGVGDRAYLQQAMDWFARHKDFIVVYTSNGMKWCLENIDTSHGDVVFA3	384
QY	249	NGTEGSPAKDPALLTCQNHTFMITIGIFGIWAAYLAGSDTIYLANYTLPDSPFLVKFKPZA	308
Dd	285	NGOEGTPGPKDPALLTCQNHTFMITIGIFGWAAVLAGDIVILANFTLPDSEFLUKIFPEA	344
QY	309	AFLPEWVGIPADLSPLLKALTACPRSHFLH	339
Dd	345	AFLPEWGINADLSPLIQAFDPKPKPSLFRL	375
RESULT 7			
ID	FUT1_RAT	STANDARD;	PRT; 376 AA.
IC	QI0980;		
DT	01-OCT-1996 (Rel. 38, Created)		
DT	15-JUL-1999 (Rel. 38, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2) _{FT} 1)		
DE	(Fucosyltransferase 1).		
GN	FUT1 OR FTA.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;		
OX	NCBI_TaxID=101116;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Colon cancer;		
RC	Scejlma M., Wang B., Koda Y., Kimura H.;		
RT	"Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-		
RT	fucosyltransferase genes.";		
RL	Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.		
RL	[2]		
RP	SEQUENCE OF 169-310 FROM N.A.		
RC	STRAIN=B0IX;		
KX	MEDLINE=94280382; PubMed=8010942;		
RT	Piau J.-P., Labarrière N., Dabouis G., Denis M.G.;		
RT	"Evidence for two distinct alpha(1,2)-fucosyltransferase genes		
RT	differentially expressed throughout the rat colon.";		
CC	Biochem. J. 300:623-626(1994).		
CC	-i- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +		
CC	alpha-L-fucosyl-1,2-beta-D-galactosyl-R.		
CC	-i- PATHWAY: Glycosylation.		
CC	-i- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND		
CC	FORM IN TRANS CITERNAE OF GOLGI (BY SIMILARITY).		
CC	-i- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE		
CC	GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A		
CC	TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF		
CC	MESOENTERAL OR ENDODERMAL ORIGIN RESPECTIVELY.		
CC	-i- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.		

CC	This SWISS-PROT entry is copyright. It is produced through a collabora-		
CC	tion between the Swiss Institute of Bioinformatics and the EMBL Outstat-		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no		
CC	modified and this statement is not removed. Usage by and for commer-		
CC	entities requires a license agreement (See http://www.isb-sib.ch/annou-		
CC	or send an email to licenses@isb-sib.ch).		

DR	EMBL; AB015637; BAA31130.1; -;		
DR	EMBL; AB0006137; BAA21741.1; -;		
DR	EMBL; L26009; AAB41514.1; -;		
DR	InterPro; IPR002516; GT 11.		
DR	Pfam; PF01531; Glyco_transf.11; 1.		
KW	Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;		
KW	Signal-anchor; Golgi stack.		
DOMAIN	1 12		
FT	TRANSMEM 13 29		
FT	CYTOPLASMIC (POTENTIAL).		
FT	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)		
FT	(POTENTIAL).		

RX MEDLINE=97240210; PubMed=9122901;
 RA Wagner F.P., Flegel W.A.;
 RT "Polymorphism of the h allele and the population frequency of
 RL sporadic nonfunctional alleles.";
 RN Transfusion 37:284-290(1997).
 [3]
 RP VARIANT HIS-164.
 RX MEDLINE=94286534; PubMed=7912436;
 RA Kelly R.J., Ernst L.K., Larsen R.D., Bryant J.G., Robinson J.S.,
 RA Lowe J.B.;
 RT "Molecular basis for H blood group deficiency in Bombay (Oh) and
 RL para-Bombay individuals.";
 RN Proc. Natl. Acad. Sci. U.S.A. 91:5843-5847(1994).
 [4]
 RP VARIANT ARG-242.
 RX MEDLINE=97445117; PubMed=9299444;
 RA Koda Y., Soejima M., Johnson P.H., Smart E., Kimura H.;
 RT "Missense mutation of FUT1 and deletion of FUT2 are responsible for
 RL Indian Bombay phenotype at ABO blood group system.";
 RN Biochem. Biophys. Res. Commun. 238:21-25(1997).
 CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
 CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
 CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
 CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
 CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
 CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI.
 CC -!- POLYMORPHISM: NONFUNCTIONAL MUTANT OF FUT1 ARE THE CAUSE OF THE H-
 CC BOMBAY BLOOD GROUP.
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
 CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
 CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
 CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M35531; AAA52639.1; -;
 DR EMBL; Z69587; CAA93435.1; -;
 DR PIR; A36047; A36047.
 DR Genew; HGNC:4012; FUT1.
 DR MIN; 211100; -;
 DR InterPro; IPR002516; GP_11.
 DR Pfam; PF01531; Glyco_transf_11; 1.
 DR Trnasf; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Transferase; Glycosyltransferase; Blood group antigen.
 KW Signal-anchor; Golgi stack; Polymorphism; Cytophasmic (POTENTIAL).
 KW DOKAIN 1 8
 FT TRANSFER 9 25
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT LUMENAL, CATALYTIC (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 327 327
 FT CARBOHYD 154 154
 FT VARIANT 154 154
 FT VARIANT 164 164
 FT VARIANT 164 164
 FT VARIANT 171 171
 FT VARIANT 171 171
 FT VARIANT 242 242
 FT VARIANT 242 242
 FT VARIANT 259 259
 FT VARIANT 259 259
 FT VARIANT 315 315
 FT VARIANT 315 315
 FT VARIANT 349 349
 FT VARIANT 349 349

FT SEQUENCE 365 AA; 41251 NW; 4F4442EC375C9D9E CRC64;
 SQ /FTIQ=VAR_003421.
 Query Match 55.5%; Score 1051.5; DB 1; Length 365;
 Best Local Similarity 63.3%; Pred. No. 1.6e-81;
 Matches 195; Conservative 37; Mismatches 73; Indels 3; Gaps 2;
 QY 25 GNTESPERRDSEQHGNGELRGMFTINSIGRLGNQMGAYATLAFALARMNGRLAFTPASNH 84
 Db 59 GTAMPNASSSCPOH-PASLSGTWTVPYNGRFGNQGQYATLLAQLNGRRAFILPAKH 117
 QY 85 NALAPIFRISLPVLHSDTAKKIPQNYHLNDWMEERYHIPGHVFVFTGYPGCSWTFYHHL 144
 Db 113 AALAPVFRITLPVLAPEVDSRTPWRELQLDWMESEYADLRDPFLKLSGFPSCSWTFEHL 177
 QY 145 RPEILKEFTLHDHVREREAQAFLRGRV--NGSOPSFVGVHVHVRGDDYVHVNVNWKGVVA 202
 Db 173 RQIIRKEFTLHDHVREREAQAFLRGRV--NGSOPSFVGVHVHVRGDDYVHVNVNWKGVVG 237
 QY 203 DRGYLEKALDMFRARYSSPVFVVTNSNGMAWCRENINASRGDVFVAGNIEGSPAKDFALL 262
 Db 233 DSAYLRQAMDFRAREAPVFFVVTNSNGMECKENIDTSQGDVTFAGDQGEATFWKDFALL 297
 QY 263 TQCNHITMTIGTIGWAAYLAGGDIYIANYTLPPSPFLKVKFPEAAFLPEWVGIPADLS 322
 Db 298 TQCNHITMTIGTIGWAAYLAGGDIYIANYTLPPSPFLKVKFPEAAFLPEWVGIPADLS 357
 QY 323 PLLKALTP 330
 Db 358 PLWTLAKP 365
 RESULT 10
 ID FUT1_PIG STANDARD; PRT; 365 AA.
 AC Q29043; O19101;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
 DE D-galactoside 2-alpha-L-fucosyltransferase 1 (Alpha(1,2)Ft 1)
 DE (Fucosyltransferase 1).
 GN Fut1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96217559; PubMed=8613146;
 RA Cohnsey S., Moutouris E., McKenzie I.F., Sandrin M.S.;
 RT "Molecular cloning of the gene coding for pig alpha1-->2
 RT fucosyltransferase";
 RL Immunogenetics 44:76-79(1996).
 RN [2]
 RP SEQUENCE FROM N.A. AND VARIANTS THR-103 AND GLN-286.
 RX MEDLINE=97468270; PubMed=9321466;
 RA Meijerink E., Fries R., Voegelé P., Masabanda J., Wigger G.,
 RA Stricker C., Neuschwander S., Bertschinger H.U., Stranzinger G.;
 RT "Two alpha(1,2) fucosyltransferase genes on porcine chromosome eq1
 RT are closely linked to the blood group inhibitor (S) and Escherichia
 RT coli F18 receptor (ECF18R) loci.";
 RL Mamm. Genome 8:736-741(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Small intestine;
 RA Meijerink E., Neuschwander S., Fries R., Dinter A.,
 RA Bertschinger H.U., Stranzinger G., Voegelé P.;
 RT "Expression and activity of porcine alpha(1,2)fucosyltransferases
 RT determine erythrocyte antigen precursor O status and susceptibility to
 RT Escherichia coli F18 colonization in the small intestine.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +

alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
 -!- PATHWAY: Glycosylation.
 -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 -!- FORM IN TRANS CISTERNAE OF GOLGI.
 -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
 GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
 TISSUE-SPECIFIC MANNER.
 -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: L50534; AAB02984.1; -.
 EMBL: U70883; AAB81884.1; -.
 EMBL: AF136896; AAF59833.1; -.
 InterPro: IPR002516; GT_11.
 Pfam: PF01531; Glyco_transf_11; 1.
 Transferrase: Glycosyltransferase; Glycoprotein; Transmembrane;
 Signal-anchor; Golgi stack; Polymorphism.
 KW DOMAIN 1 8
 TRANSMEM 9 28
 DOMAIN 29 365
 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 65 65
 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301
 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 327 327
 A -> T.
 FT VARIANT 103 103
 R -> Q.
 FT CONFLICT 77 77
 L -> F (IN REF. 1).
 SQ SEQUENCE 365 AA; 41106 MW; DAFCE77E89A29D75 CRC64;

Query Match 55.4%; Score 1050.5; DB 1; Length 365;
 Best Local Similarity 62.9%; Pred. No. 1.9e-81;
 Matches 195; Conservative 38; Mismatches 74; Indels 3; Gaps 2;
 QY 17 PWTOMSSNGTSPESPMRDSQHGNGELRGMFTINSIGRLGNQMGEBYATLFAALRMNGRL 76
 DB 51 PVAICLAGTFVHPNADSCPKH-PASLSGTWTITPDGRFGNQMGQYATLLAALQNGRQ 109
 QY 77 AFIPASMHNALAPIRISLPLVLSHTAKKIPQNYHLNDWMEERYHPHFVRETGYPC 136
 DB 110 AFIQAPAMHVALPVPRTILPLVLAPEVDHAPWKELEHDMSEDYAHLKEPWLKLTGFPC 169
 QY 137 SWTFYHHLRPEILKFTLHDHVBREAQAEFLRLV--NGSQPSTFGVGVHVRGDIYVHVP 194
 DB 170 SWTFYHHLRPEILKFTLHDHVBREAQAEFLRLV--NGSQPSTFGVGVHVRGDIYVHVP 229
 QY 195 NWKGVVADRGVLEKALDMFRARYSSPVFVTSNGMAWCRNENINASRGDGVVAGNGIEGS 254
 DB 230 KRWKGVGVGAYLQAMDMFRARYEPAPVFTVTSNGMEWCKRKNIDTSRGDIVIFAGDGRRA 289
 QY 255 PAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANVTLTPDSPPLKVKFPEAAFLPEW 314
 DB 290 PAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANVTLTPDSPPLKVKFPEAAFLPEW 349
 QY 315 VGIPADLSPL 324
 DB 350 VGINADLSPL 359

RESULT 11
 FUT2_RAT
 ID FUT2_RAT STANDARD; PRT: 159 AA.
 AC Q10984;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.59) (Secretor blood
 group alpha-2-fucosyltransferase) (GDP-L-fucose:beta-D-

DE galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)FT 2)
 DE (Fucosyltransferase 2) (Fragment).
 GN FUT2 OR SEC1 OR ETB.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDIX;
 RX MEDLINE=94280382; PubMed=8010942;
 RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.:
 RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
 RT differentially expressed throughout the rat colon.";
 RL Biochem. J. 300:623-626(1994).
 CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE
 CC FUC-ALPHA(1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN
 CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED
 CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE
 CC THE SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
 CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
 CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
 CC TISSUE-SPECIFIC MANNER.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: L26010; AAB41515.1; -.
 InterPro: IPR002516; GT_11.
 Pfam: PF01531; Glyco_transf_11; 1.
 Transferrase: Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Golgi stack.
 FT NON_TER 1 1
 CARBOHYD 98 98
 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 159 AA; 17250 MW; 14DECEB7C2E6384A CRC64;
 Query Match 44.6%; Score 846; DB 1; Length 159;
 Best Local Similarity 99.4%; Pred. No. 1.4e-64;
 Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 195 NWKGVVADRGVLEKALDMFRARYSSPVFVTSNGMAWCRNENINASRGDGVVAGNGIEGS 254
 DB 1 NWKGVVADRGVLEKALDMFRARYSSPVFVTSNGMAWCRNENINASRGDGVVAGNGIEGS 60
 QY 255 PAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANVTLTPDSPPLKVKFPEAAFLPEW 314
 DB 61 PAKDFALLTQCNTHTIMTIGTFGIWAAYLAGGDTIYLANVTLTPDSPPLKVKFPEAAFLPEW 120
 QY 315 VGIPADLSPL 324
 DB 121 VGIPADLSPL 359

RESULT 12
 YKQ7_CABEL
 ID YKQ7_CABEL STANDARD; PRT: 365 AA.
 AC P34302;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative glycosyl transferase C06E1.7 in chromosome III (EC 2.4.1.59).
 GN C06E1.7.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94150718; PubMed=7906398;
 RX Watson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laistner N.,
 RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons C., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS.
 RA Waterston R.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2002) to the GLYCOSYLTRANSFERASE FAMILY 11.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL: L16559; AAA27932.2;
 DR WormPep: C05E1.7; CE30483;
 DR InterPro: IPR002516; GT11;
 DR Pfam: PF01531; Glyco_transf_11;
 KW Hypothetical protein; Transferase; Glycosyltransferase.
 SQ SEQUENCE 365 AA; 41991 MW; B5FBCA363F31977F CRC64;
 Query Match 8.9%; Score 168; DB 1; Length 365;
 Best Local Similarity 22.0%; Pred. No. 7.3e-07;
 Matches 71; Conservative 48; Mismatches 115; Indels 88; Gaps 16;
 QY 52 SIGRLGNMGGEYATLFAARMNGR--LAFIPASMHNALAPTRISLPVLHSD----- 101
 DB 70 STARLANHIFELVSIVGNAKSLNKPALFVDESKYNLLITGVKVLPLGLDFEQIFPV 129
 QY 102 --TAKKP-----WQNYHLNDWMEERYRHIPGHVFRTGYPCSWTFYHHLRPEILKEF 152
 DB 130 HNKATKVPLEKCKCIFDPDKFNISSELYLHTGHFYQ-----SWKYF----- 172
 QY 153 TLHDHVRREAQAFRLGL-----RVNGSOPSTFVG---VHVRGDYV---HVMPNVKGVVA 202
 DB 173 ---DHYEKVQSVFKPALDFSPSPNSDSNFISRICTHIRRTDFDGGHSSNV----- 223
 QY 203 DRGYLEKALDMFRAR-----YSSPVFVVTSGMCAWENINASRGD 243
 DB 224 --STPKPALEFIKERQKDNKMLTVIMGDDPDPEAKMF-----EGTVRAKKEAKIETT 277
 QY 244 VFRAGNGIEGSPAKDFALL--TQCNIIMTY--GTFGIWAAYLAGSDTITLANY--TLDPSP 299
 DB 278 KYFVS---ENTPQDQLAYSHYSCDATTLTAPSSFTGWLGLYLSKGQAVYYQDIRSTNDV 334
 QY 300 FLK-VFKPEAPAFPEWICIPAD 320
 DB 335 YKKGVLDPDDFFVPVSWTSIMLD 356
 RESULT 13

VGLD_HSVK *
 ID VGLD_HSVK STANDARD; PRT; 442 AA.
 AC P22484;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein D precursor (glycoprotein 17/18).
 GN GD OR GP17/18 OR 72.
 OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OC NCBI_TaxID=10329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91082407; PubMed=1845821;
 RA Flowers C.C., Eastman E.M., O'Callaghan D.J.;
 RT "Sequence analysis of a glycoprotein D gene homolog within the unique
 RT short segment of the EHV-1 genome.";
 RL Virology 180:173-184(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92263759; PubMed=1316673;
 RA Colle C.F. III, Flowers C.C., O'Callaghan D.J.;
 RT "Open reading frames encoding a protein kinase, homolog of
 RT glycoprotein gx of pseudorabies virus, and a novel glycoprotein map
 RT within the unique short segment of equine herpesvirus type 1.";
 RL Virology 188:545-557(1992).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL: M62923; AAA46081.1;
 DR EMBL: M86931; -- NOT_ANNOTATED_CDS.
 DR EMBL: M87497; AAA46073.1; ALT_INIT.
 DR PIR: A38518; VGBEEA.
 DR InterPro: IPR002896; Herpes_glycop_D.
 DR Pfam: PF01537; Herpes_glycop_D; 1.
 KW Glycoprotein; Signal; Transmembrane.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 442 GLYCOPROTEIN D.
 FT DOMAIN 20 405 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 406 422 POTENTIAL.
 FT DOMAIN 423 442 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 442 AA; 49908 MW; 323CDCA9C9762F05 CRC64;
 Query Match 4.8%; Score 91; DB 1; Length 442;
 Best Local Similarity 22.7%; Pred. No. 2.9;
 Matches 54; Conservative 31; Mismatches 103; Indels 50; Gaps 10;
 QY 1 LQORIVKIQ-----PLSEKELPMTTOMSSCNTSPENRRSEGHGNGELRGMTI 50
 DB 234 LYRRVIEIDGGRIYDFSVTPSERCPIAFELNFGN---PORCKTPEQYSRGE---VTR 287
 QY 51 NSIGRLGNMGGEYAT--LFALARMNGRLAFIPASMHNALAPTRISLPVLHSDTAKKIPW 108
 DB 288 RFLGLENFPQGEHMTWKVFWFYDGGNL--PVQFYEAQ-----FAPVPZP 331
 QY 109 QNYHLNDWMEERYRH-----IPGHFVRFRTGYPCSWTFYHHLRPEILKEFTLHDHVRQAQ 163
 DB 332 DNHGPGDSVESEITQNKTDPKPGQADPKPNQPFKWPISKHLVPLR-----DEVDEVIE 384
 QY 164 AFLRLRVNGSQPSTFVGHVHVRGDYVHVMPNVKGVVADVADRGYLEKALDMFRARVSSP 222

Db 385 PVTPEPKTSKN-STFVGISVGLGIAGLVGVILYVCLRR---KKELKVCTERLDSP 438

RESULT 14

FMA3_FORGI

AC Q51826; STANDARD; PRT; 353 AA.

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Major fimbrial subunit protein, type III precursor (Fimbrillin)

DE (Fimbrillin)

GN FIMA.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;

OC Porphyromonadaceae; Porphyromonas.

OX NCBI_TaxID=837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=6/26;

RX MEDLINE=94071950; PubMed=7902712;

RA Fujiwara T., Morishima S., Takahashi I., Hamada S.;

RT "Molecular Cloning and sequencing of the fimbrillin gene of

RT Porphyromonas gingivalis strains and characterization of recombinant

RT proteins.";

RL Biochem. Biophys. Res. Commun. 197:241-247(1993).

RN [2]

RP FUNCTION, AND CLASSIFICATION INTO TYPES.

RX MEDLINE=21614934; PubMed=11748193;

RA Nakagawa I., Amano A., Kuboniwa M., Nakamura T., Kawabata S.,

RA Hamada S.;

RT "Functional differences among FimA variants of Porphyromonas

RT gingivalis and their effects on adhesion to and invasion of human

RT epithelial cells.";

RL Infect. Immun. 70:277-285(2002).

CC -!- FUNCTION: Fimbrillin is the structural subunit of the fimbriae,

CC that are filamentous appendages on the cell surface. Fimbriae of

CC P.gingivalis are recognized as a major virulence factor as they

CC mediate cell adhesion and play an important role in invasion of

CC periodontal tissues.

CC -!- SUBCELLULAR LOCATION: Fimbria.

CC -!- SIMILARITY: BELONGS TO THE P.GINGIVALIS FIMBRILLIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D17801; BAA04627.1; -.

KW Fimbria; Virulence.

FT PROPEP 1 10 BY SIMILARITY.

FT CHAIN 11 353 MAJOR FIMBRILLIN SUBUNIT PROTEIN, TYPE III.

SO SEQUENCE 353 AA; 38024 MW; 7FBE4FBF427EA2AB CRC64;

Query Match 4.7%; Score 89; DB 1; Length 353;

Best Local Similarity 22.1%; Pred. No. 3.3;

Matches 64; Conservative 38; Mismatches 110; Indels 78; Gaps 12;

Qy 18 MTTOMSSNTSPEMRDSEQ-----HGNGELRGMTFINSIGRLGNQMG---EYATLFA 68

Db 80 LTTTELGNOEAAGLIMTAEPVEVLVAGN-----NYGYDGSQGGNOISGCTPLE 130

Qy 69 LARMNGRLAF--IPASHMNALAPIRISLPVLHSDTAKK----- 105

Db 131 IKRVHARIAFTKIEVMSQSVANKNFAPENIYALVAKKSNLFGASLANSDDAYLTGSL 190

Qy 106 -----IPQWYHLDNDWEERYRHI-----PGHFVFTGYPCSWTYYH--LRPEIL- 149

Db 191 ITFNGAYSANYTHVDNLGROYTEGANTVTPKGF-----YVLESTYAQNAGLRPTILC 245

Qy 150 --KBETLHDH---REEAQAFRLGLRYNGSQSPSF--VGHVHRRGDYVHVPNPVWKGVV 201

Db 246 VKGKLTKHGDGTALSSEMTAAFNAGWIVANNDPTTYYPVLVNFESNNYTYTGEAVEKGKI 305

Qy 202 ADRGYLEKALDMFRARYSSPVFVTSNGMNCREINASRGDVFVAGNGI 251

Db 306 VRNHKFDINLTITGFTNNPENPIT-----ESANLNVCVVAARKGV 347

RESULT 15

CDH_PHACH

ID CDH_PHACH STANDARD; PRT; 773 AA.

AC Q01738; Q00047;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cellobiose dehydrogenase precursor (EC 1.1.5.1) (CDH) (cellobiose-

DE quinone oxidoreductase).

GN CDH-1 AND CDH-2.

OS Phanerochaete chrysosporium.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Corticiaceae; Phanerochaete.

OX NCBI_TaxID=5306;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97077226; PubMed=8919793;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cloning of a cDNA encoding cellobiose dehydrogenase, a

RT hemoflavoenzyme from Phanerochaete chrysosporium.";

RL App. Environ. Microbiol. 62:1329-1335(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OGC101;

RX MEDLINE=97175414; PubMed=90233960;

RA Li B., Nagalla S.R., Renganathan V.;

RT "Cellobiose dehydrogenase from Phanerochaete chrysosporium is encoded

RT by two allelic variants.";

RL App. Environ. Microbiol. 63:796-799(1997).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 19-208.

RX MEDLINE=20135694; PubMed=10673428;

RA Hallberg B.M., Bergfors T., Boeckbro K., Pettersson G., Henriksson G.,

RA Divne C.;

RT "A new scaffold for binding haem in the cytochrome domain of the

RT extracellular flavocytochrome cellobiose dehydrogenase.";

RL Structure 8:79-88(2000).

CC -!- FUNCTION: Degrades BOTH LIGNIN AND CELLULOSE. OXIDIZES CELLOBIOSE

CC TO CELLOBIONOLACTONE.

CC -!- CATALYTIC ACTIVITY: Cellobiose + a quinone = cellobiono-1,5-

CC lactone + a phenol.

CC -!- COFACTOR: ONE FAD AND ONE HEME B.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GMC

CC OXIDOREDUCTASES FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U46081; AAC49277.1; -.

DR EMBL; U65888; AB61455.1; -.

DR EMBL; U50409; AAB92262.1; -.

DR PDB; 1D7B; 18-OCT-99.

DR PDB; 1D7C; 18-OCT-99.

DR PDB; 1D7D; 18-OCT-99.

DR InterPro; IPR000172; GMC_oxred.

DR InterPro; IPR001100; Pyr_redox.

DR Pfam; PF00732; GMC_oxred; 1.

Search completed: May 27, 2003, 15:07:14
Job time : 11.3317 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:52:27 ; Search time 43.6293 Seconds

(without alignments)

1050.630 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVQMPSPFMAHRLTFV.....AAFLPEWTGIAADLSPLLKH 344

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	344	21	Human Sec2 catalyt
2	1832.5	99.1	343	19	Human Sec2 protein
3	1528	82.6	340	19	Porcine secretor t
4	1528	82.6	340	19	Pig secretor. Sus
5	1475	79.7	380	21	Rat hepatoma H35 c
6	1475	79.7	380	23	Rat hepatoma H35 c
7	1376	74.4	353	21	Rat hepatoma H35 c
8	1376	74.4	353	23	Rat hepatoma H35 c
9	1081.5	58.5	365	20	Swine alpha(1,2) f
10	1081.5	58.5	365	20	Swine alpha-1,2-fu

11	1080.5	58.4	365	12	AA13751
12	1080.5	58.4	365	16	AA180154
13	1080.5	58.4	365	16	AA180154
14	1080.5	58.4	365	16	AA180154
15	1080.5	58.4	365	17	AA180154
16	1080.5	58.4	365	18	AA180154
17	1080.5	58.4	365	18	AA180154
18	1080.5	58.4	365	18	AA180154
19	1080.5	58.4	365	19	AA180154
20	1080.5	58.4	365	21	AA180154
21	1080.5	58.4	365	21	AA180154
22	1080.5	58.4	365	23	AA180154
23	1080.5	58.4	365	23	AA180154
24	1080.5	58.4	365	20	AA180154
25	1080.5	58.4	365	21	AA180154
26	1080.5	58.4	365	21	AA180154
27	1080.5	58.4	365	21	AA180154
28	1080.5	58.4	365	21	AA180154
29	1080.5	58.4	365	21	AA180154
30	1080.5	58.4	365	21	AA180154
31	1080.5	58.4	365	21	AA180154
32	1080.5	58.4	365	21	AA180154
33	1080.5	58.4	365	21	AA180154
34	1080.5	58.4	365	21	AA180154
35	1080.5	58.4	365	21	AA180154
36	1080.5	58.4	365	21	AA180154
37	1080.5	58.4	365	21	AA180154
38	1080.5	58.4	365	21	AA180154
39	1080.5	58.4	365	21	AA180154
40	1080.5	58.4	365	21	AA180154
41	1080.5	58.4	365	21	AA180154
42	1080.5	58.4	365	21	AA180154
43	1080.5	58.4	365	21	AA180154
44	1080.5	58.4	365	21	AA180154
45	1080.5	58.4	365	21	AA180154

ALIGNMENTS

RESULT 1
ID AAB36106 standard; Protein; 344 AA.
XX AAB36106;
AC AAB36106;

DT 19-FEB-2001 (first entry)

DE Human Sec2 catalytic domain.

XX Human; Sec2; alphas-2fucosyltransferase; cytosolic;

KW neuroprotective; nootropic; gene therapy; Fucal-2Gal-beta-1-3GalNAc;

KW immunotherapy; immunosuppression; cancer; neurological disease;

KW small cell lung carcinoma.

XX Homo sapiens.

OS WO2000064464-A1.

PN 02-NOV-2000.

PD 23-APR-1999; 99WO-US07384.

XX 23-APR-1999; 99WO-US07384.

XX (PACI-) PACIFIC NORTHWEST CANCER FOUND.

PI Holmes EH, Sherwood AL;

PI WPI; 2000-687262/57.

DR New rat ganglioside GM1-specific alpha-2fucosyltransferase, useful for

XX preparation of fucosyl GM1 which is useful as a nutritional composition

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT

PT or immunotherapeutic for cancer and neurological diseases -

PS Example; Fig 3B; 91pp; English.

XX The present sequence is given in a specification relating to a rat
CC ganglioside GM1-specific alpha1-2-fucosyltransferase protein. The protein
CC or its cellular fraction is useful for synthesis of a molecule comprising
CC Fucal-2-galactose-3-galNAc, a glycolipid, glycoprotein, glycolipoprotein
CC or a free oligosaccharide comprising Fucal-2-galactose-3-galNAc.
CC The method involves contacting alpha1-2-fucosyltransferase with GDP-fucose
CC and a molecule or glycolipid, glycoprotein, glycolipoprotein or
CC oligosaccharide having a terminal galactose-3-galNAc group. It is also
CC useful for synthesis of fucosyl-GM1 by contacting the protein with
CC GDP-fucose and ganglioside GM1. The obtained glycoproteins,
CC glycolipoproteins, glycolipids and oligosaccharides are useful as
CC nutritional compositions and fucosyl-GM1 is useful for inducing an
CC immunotherapeutic or immunosuppressive action against cancer,
CC neurological disease or small cell lung carcinoma.

XX Sequence 344 AA;

Query Match 100.0%; Score 1850; DB 21; Length 344;
Best Local Similarity 100.0%; Pred. No. 2.5e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLVQMPFSPMAHFLFVFTVSTIFHVQQLAKIQAMWELPVOIPVLASTSKALGPSQL 60
DB 1 MLVQMPFSPMAHFLFVFTVSTIFHVQQLAKIQAMWELPVOIPVLASTSKALGPSQL 60
QY 61 RGMWTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSSTLAPFIRITLPVLHSATAS 120
DB 61 RGMWTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSSTLAPFIRITLPVLHSATAS 120
QY 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180
DB 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180
QY 181 KFLRGQVNSRPGTFVGVVHVRGDIYVHVPKWKGVVADRRYLOQALDWFPRARYSSLIIF 240
DB 181 KFLRGQVNSRPGTFVGVVHVRGDIYVHVPKWKGVVADRRYLOQALDWFPRARYSSLIIF 240
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTIGTGIWAAYILT 300
DB 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTIGTGIWAAYILT 300
QY 301 GGDITILANYTLPPDSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344
DB 301 GGDITILANYTLPPDSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344

RESULT 2
AAW69332
ID AAW69332 standard; Protein; 343 AA.

XX AAW69332;

XX 20-NOV-1998 (first entry)

DE Human Sec2 protein sequence.

XX Sec2; alpha(1,2) fucosyltransferase; H blood group; secretor genotyping;
KW GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase; human;
KW FUT2; nonsecretor genotyping.

XX Homo sapiens.

XX US5807732-A.

XX 15-SEP-1998.

XX 28-FEB-1995; 95US-0395800.

XX 28-FEB-1995; 95US-0395800.

XX (GIOR/) GIORGI D.
PA (KELL/) KELLY R.J.
PA (LENN/) LENNON G.
PA (LOWE/) LOWE J.B.
PA (ROUQ/) ROUQUIER S.

XX Giorgi D, Kelly RJ, Lennon G, Lowe JB, Rouquier S;
PI WPI: 1998-520127/44.
XX DR N-PSDB; AAV58323.

XX DNA encoding fucosyltransferase enzyme - useful for producing
PT recombinant enzyme and genotyping person as secretor or nonsecretor
PT Disclosure; Column 45-50; 55pp; English.

XX This sequence is the human Sec2 protein of the invention. The DNA
CC encodes a alpha(1,2) fucosyltransferase and is the Secretor
CC alpha(1,2)fucosyltransferase locus, that cross hybridises with the
CC H blood group alpha(1,2)fucosyltransferase gene. The DNA is useful for
CC producing a recombinant human GDP-L-fucose:beta-D-galactoside
CC 2-alpha-L-fucosyltransferase (FUT2) which can be used for genotyping an
CC individual as a secretor or nonsecretor as it is known that nonsecretors
CC homozygous for a mutant allele of the FUT2 gene that has a stop codon in
CC the position corresponding to amino acid 143.

XX Sequence 343 AA;

Query Match 99.1%; Score 1832.5; DB 19; Length 343;
Best Local Similarity 99.7%; Pred. No. 1.8e-184;
Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVQMPFSPMAHFLFVFTVSTIFHVQQLAKIQAMWELPVOIPVLASTSKALGPSQL 60
DB 1 MLVQMPFSPMAHFLFVFTVSTIFHVQQLAKIQAMWELPVOIPVLASTSKALGPSQL 60
QY 61 RGMWTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSSTLAPFIRITLPVLHSATAS 120
DB 61 RGMWTINAGRLGNQMGYATLYALAKMNGRPAPFPAQMHSSTLAPFIRITLPVLHSATAS 120
QY 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 180
DB 121 RHPWNYHLNDWMEEEYRHIPPGEYVFTGYPSCWTFYHHLRQELQFTLHDHVRREAAQ 179
QY 181 KFLRGQVNSRPGTFVGVVHVRGDIYVHVPKWKGVVADRRYLOQALDWFPRARYSSLIIF 240
DB 180 KFLRGQVNSRPGTFVGVVHVRGDIYVHVPKWKGVVADRRYLOQALDWFPRARYSSLIIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTIGTGIWAAYILT 300
DB 240 VVTSNGMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTMTIGTGIWAAYILT 299
QY 301 GGDITILANYTLPPDSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 344
DB 300 GGDITILANYTLPPDSPFLKIFKPEAAFLPEWTGIAADLSPLLKH 343

RESULT 3

AAW37855

ID AAW37855 standard; Protein; 340 AA.

XX AAW37855;

XX 28-AUG-1998 (first entry)

XX Porcine secretor transferase (FUT2).

XX Secretor; glycosyltransferase; FUT2; pig; epitope; antigen;
KW transgenic animal; xenotransplantation; organ transplant.
XX Sus scrofa.

```

FH Key Location/Qualifiers
FT Domain 1..4
FT . /note= "N-terminal cytoplasmic tail"
FT Domain 5..26
FT . /note= "transmembrane domain"
FT Domain 27..340
FT . /note= "C-terminal domain"
FT Modified-site 185..187
FT . /note= "Asn is N-glycosylated"
FT Modified-site 251..253
FT . /note= "Asn is N-glycosylated"
FT Modified-site 279..281
FT . /note= "Asn is N-glycosylated"
FT Modified-site 305..307
FT . /note= "Asn is N-glycosylated"
XX
XX WO9807837-A1.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-AU00540.
XX
XX 23-AUG-1996; 96AU-0001823.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX McKenzie IFC, Sandrin MS;
XX
XX WPI: 1998-169148/15.
XX
XX N-PSDB; AAV29003.
XX
XX Nucleic acid encoding glycosyltransferase able to compete with
XX second such enzyme - particularly used to reduce expression of
XX unwanted carbohydrate epitope(s) on tissues intended for
XX transplantation
XX
XX Claim 6; Fig 1A-B; 40pp; English.
XX
XX This polypeptide comprises porcine secretor glycosyltransferase
XX (SE or FUT2), a type II integral membrane protein has high affinity
XX for type I and type III substrates. Its amino acid sequence was
XX deduced from the nucleotide sequence of a genomic DNA clone (see
XX AAV29003) isolated from a pig liver library on the basis of homology
XX to human FUT2 cDNA. Pig FUT2 shows 83.2% identity with human FUT2.
XX Expression of a glycosyltransferase, such as porcine Se, in a
XX tissue results in reduced expression of unwanted carbohydrate
XX epitopes on the tissue, especially porcine heart, liver, kidney or
XX pancreas, rendering it more suitable for transplantation, i.e. less
XX immunogenic and of increased immunological acceptability. A
XX claimed method of producing a cell from a donor species that is
XX immunologically acceptable to a recipient species involves reducing
XX levels of carbohydrate on the donor cell that causes it to be
XX recognised as non-self by the recipient by expressing a nucleic
XX acid for a glycosyltransferase such as porcine Se in the cell.
XX
XX Sequence 340 AA;
XX
XX Query Match 82.6%; Score 1528; DB 19; Length 340;
XX Best Local Similarity 82.0%; Pred. No. 2.3e-152;
XX Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;
XX
QY 1 MLVQMPSFPMHAFILFVFTYSTFHVQORLAKTCAMWELPVQIPVLASTSKALGPSOL 60
Db || : || || || || || || || || || || || || || || || || || || ||
QY 61 RGMWITNATIGLNGOMGEVATILYALAKMNGRAFPAPQMHSTLAFITRILPVLASATAS 120
Db : || || || || || || || || || || || || || || || || || || || ||
QY 58 KGMWITNATIGLNGOMGEVATILYALAKMNGRAFPAPQMHSTLAFITRILPVLASTAR 117
Db : || || || || || || || || || || || || || || || || || || || ||
QY 121 RIPWQNYHLNDWMEERYHIPPGEYVRFTGPCSWTFYHHLRQETIQOFTLHDHVREPAQ 180
Db || || || || || || || || || || || || || || || || || || || ||
QY 116 RIPWQNYHLNDWMEERYHIPPGEYVRFTGPCSWTFYHHLRQETIQOFTLHDHVREPAQ 176
Db || || || || || || || || || || || || || || || || || || || ||

```

```

QY 181 KFLRGLQVNGSRPGTFVGVHVRGDDYVHVMPKVMKGWADRRYVQQALDWFRRARYSSLIF 240
Db || || || || || || || || || || || || || || || || || || || ||
QY 177 DELRGLRVNGSRPSTYGVGVHVRGDDYVHVMPKVMKGWADRRYVLEQALDWFRRARYSPVF 236
Db || || || || || || || || || || || || || || || || || || || ||
QY 241 VVTSNGMWCNRINIDFSGDVVFGAGDIESSPAKDFALLTQCQNHITWTICTGFIWAAYLT 300
Db || || || || || || || || || || || || || || || || || || || ||
QY 237 VVSSNGMWCNRININASGDVWFAGDIESSPAKDFALLTQCQNHITWTICTGFIWAAYLA 296
Db || || || || || || || || || || || || || || || || || || || ||
QY 301 GGDTIYLANITLPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 344
Db || || || || || || || || || || || || || || || || || || || ||
QY 297 GGDTIYLANITLPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 340
Db || || || || || || || || || || || || || || || || || || || ||
XX
XX RESULT 4
XX AAW53101
XX ID AAW53101 standard; Protein; 340 AA.
XX
XX AC AAW53101;
XX
XX XX 08-JUL-1998 (first entry)
XX
XX DE Pig secretor.
XX
XX KW Pig; secretor; chimeric; glycosyltransferase; gene therapy;
XX KW transplantation.
XX
XX OS Sus scrofa.
XX
XX PN WO9805768-A1.
XX
XX PD 12-FEB-1998.
XX
XX PF 01-AUG-1997; 97WO-AU00492.
XX
XX PR 21-AUG-1996; 96US-0024279.
XX PR 02-AUG-1996; 96AU-0001402.
XX
XX PA (AUST-) AUSTIN RES INST.
XX
XX PI McKenzie IFC, Sandrin MS;
XX
XX XX WPI: 1998-159170/14.
XX DR N-PSDB; AAV21639.
XX
XX Nucleic acids encoding chimeric glycosyltransferases - used for
XX altering carbohydrate levels on the surface of cells, useful in gene
XX therapy and transplantation
XX
XX Example 2; Fig 6; 51pp; English.
XX
XX The present sequence represents pig secretor used in an example of the
XX present invention. The present invention describes nucleic acids (NA)
XX encoding a chimeric glycosyltransferase. The NAs comprise a catalytic
XX domain of a first glycosyltransferase (GT) and a localisation signal of
XX a second GT, whereby when the NA is expressed in a cell and where the
XX chimeric enzyme is located in an area of the cell where it is able to
XX compete for substrate with a second GT, resulting in reduced levels of
XX a product from the second GT. The NAs can be used to produce cells and
XX organs with desired glycosylation patterns. Products and methods of the
XX present invention can be used to reduce the levels of undesirable
XX epitopes in cells, tissues or organs which may be used in
XX transplantation or gene therapy.
XX
XX Sequence 340 AA;
XX
XX Query Match 82.6%; Score 1528; DB 19; Length 340;
XX Best Local Similarity 82.0%; Pred. No. 2.3e-152;
XX Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;
XX
QY 1 MLVQMPSFPMHAFILFVFTYSTFHVQORLAKTCAMWELPVQIPVLASTSKALGPSOL 60
Db || : || || || || || || || || || || || || || || || || || || ||
QY 1 MLVQMPSFPMHAFILFVFTYSTFHVQORLAKTCAMWELPVQIPVLASTSKALGPSOL 57
Db || : || || || || || || || || || || || || || || || || || || ||

```


QY 42 PQVIPVLAST-----SKALGPSQLRGWMTINAIQRLGNQMGVEYATLYALAKMNGRPA 93
 Db 51 PVAIFCLAGTPVHPNADSDCPKHPASFGTWIIPDGRFGNQMGQYATLLALQNGRQA 110
 QY 94 FIPAQMHSTLAPIFRITLPVLHLSATASRIPWQNYHLNDWMEEEYRHIPPQGYVRFVTPC 153
 Db 111 FIOPAMHVLAPVFRITLPVLAPEVDRHAPWRELEHDMKSEDAHLKE-PWLKLTGFPC 169
 QY 154 SWTFYHHLRQELIOEFTLHDHVRREAOKFLRGLOV--NGSRPGTFVGVHVRGQYVHVMP 211
 Db 170 SWTFYHHLRQELIOEFTLHDHVRREAOKFLRGLOV--NGSRPGTFVGVHVRGQYVHVMP 229
 QY 212 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 271
 Db 230 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 289
 QY 272 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 331
 Db 332 TGTAAADLSPL 341
 Db 350 VGINADLSPL 359

RESULT 10
 AAW30630
 ID AAW30630 standard; Protein; 365 AA.
 XX
 AC AAW30630;
 AC
 DT 01-APR-1999 (first entry)
 XX
 DE Swine alpha-1,2-fucosyltransferase 1 protein.
 XX
 KW Swine; pig; alpha-1,2-fucosyltransferase 1; FUT1; resistance;
 KW Escherichia coli; infection; oedema; postweaning diarrhoea;
 KW intestinal disorder; polymorphism.
 XX
 OS Sus scrofa.
 XX
 PN W09853102-A1.
 XX
 PD 26-NOV-1998.
 XX
 PF 20-MAY-1998; 98WO-US10318.
 XX
 PR 20-MAY-1997; 97US-0047181.
 XX
 PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.
 PA (SWIT-) SWISS FEDERAL INST TECHNOLOGY ZURICH.
 PA (USDA) US SEC OF AGRIC.
 XX
 PI Bosworth Bt, Vogeli P;
 XX
 DR WPI: 1999-059746/05.
 DR N-PSDB; AAX03811.
 XX
 PT New method of identifying swine that are resistant to intestinal
 PT colonisation by Escherichia coli - comprises use of genetic
 PT polymorphic markers, used for breeding swine resistant to
 PT Escherichia coli-related diseases
 XX
 PS Example 6; Fig 1; 35pp; English.
 XX
 CC A method has been developed for the identification of swine that are
 CC resistant to intestinal colonisation by E. coli. The method comprises
 CC determining whether a genetic polymorphism associated with resistance to
 CC colonisation is present in a swine sample, and then inferring that the
 CC swine is resistant if it is homozygous for the polymorphism. The method
 CC uses the swine alpha-1,2-fucosyltransferase (FUT1) polymorphism. The
 CC present sequence represents swine FUT1. The method enables the breeding

CC of swine that are resistant to E. coli-related diseases. This method
 CC comprises breeding swine that have a genetic polymorphism in the FUT1
 CC gene. More particularly, the identification method identifies swine that
 CC are resistant to E. coli-related intestinal disorders if, in a sample
 CC taken, the only nitrogen base at residue 307 in the FUT1 gene is adenine.
 CC Larger amplified fragments from the assay can be used for RFLP analysis,
 CC and the assay itself is used as a basis for a kit, applied to swine of
 CC any age, in detecting polymorphisms associated with E. coli F18
 CC receptors. The polymorphisms are useful in developing drugs to treat
 CC swine with E. coli-related diseases. However, a mutated form of the
 CC porcine FUT1 gene may interfere with the normal enzyme and prevent it
 CC from producing the intestinal receptor for F18. The detection of
 CC polymorphic markers in the method disclosed enables the detection and
 CC treatment of E. coli-related intestinal diseases in swine, where there
 CC has been no success using antibiotics due to unsuccessful prophylaxis.
 XX

SQ Sequence 365 AA;

Query Match 58.5%; Score 1081.5; DB 20; Length 365;
 Best Local Similarity 65.5%; Pred. No. 3.3e-105;
 Matches 203; Conservative 30; Mismatches 66; Indels 11; Gaps 3;

QY 42 PQVIPVLAST-----SKALGPSQLRGWMTINAIQRLGNQMGVEYATLYALAKMNGRPA 93
 Db 51 PVAIFCLAGTPVHPNADSDCPKHPASFGTWIIPDGRFGNQMGQYATLLALQNGRQA 110
 QY 94 FIPAQMHSTLAPIFRITLPVLHLSATASRIPWQNYHLNDWMEEEYRHIPPQGYVRFVTPC 153
 Db 111 FIOPAMHVLAPVFRITLPVLAPEVDRHAPWRELEHDMKSEDAHLKE-PWLKLTGFPC 169
 QY 154 SWTFYHHLRQELIOEFTLHDHVRREAOKFLRGLOV--NGSRPGTFVGVHVRGQYVHVMP 211
 Db 170 SWTFYHHLRQELIOEFTLHDHVRREAOKFLRGLOV--NGSRPGTFVGVHVRGQYVHVMP 229
 QY 212 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 271
 Db 230 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 289
 QY 272 KWKGVVADRRYLOQALDWFRRARYSSLIFFVVTSGMAWCRENIDTSHGVDVFAGDGEAA 331
 Db 290 PARDFALLVQCNTIMTIGTGFVWAAVLAGDITIYLANFLTSSFLKIFKPEAAFLPEW 349
 QY 332 TGTAAADLSPL 341
 Db 350 VGINADLSPL 359

RESULT 11
 AAR13751
 ID AAR13751 standard; Protein; 365 AA.
 XX
 AC AAR13751;
 XX
 DT 07-NOV-1991 (first entry)
 XX
 DE GDP-Fuc:beta-D-galactoside alpha(1,2)-fucosyltransferase.
 XX
 KW Glycosyltransferase.
 XX
 OS Homo sapiens.
 XX
 PN W09112340-A.
 XX
 PD 22-AUG-1991.
 XX
 PF 14-FEB-1991; 91WO-US00899.
 XX
 PR 12-DEC-1990; 90US-0627621.
 PR 14-FEB-1990; 90US-0479858.
 PR 14-FEB-1990; 90US-0480133.
 XX
 PA (UNMI) UNIV OF MICHIGAN.
 XX


```

OS Homo sapiens.
XX W09524495-A1.
XX PN
XX 14-SEP-1995.
XX PD
XX PF 24-JAN-1995; 95WO-US00967.
XX PR 09-MAR-1994; 94US-0208889.
XX PA (ABBO ) ABBOTT LAB.
XX PI Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
XX PI Pierce JM, Prieto PA, Smith DF;
XX DR WPI; 1995-328284/42.
XX DR N-PSDB; AAT01083.
XX PT New transgenic non-human mammal milk prods - contg. heterologous
XX PT components produced as secondary gene prods. of an heterologous gene
XX PS Example 1; Page 62-64; 83pp; English.
XX CC 2-Alpha-fucosyltransferase (AAR70422) is encoded by a cDNA clone
XX CC (AAT01083) isolated from a human epidermal carcinoma A431 library.
XX CC The enzyme can be expressed in the milk of a transgenic mammal.
XX CC esp. cow. This allows large-scale prodn. of oligosaccharides and
XX CC glycosylated proteins and lipids in the milk.
XX SQ Sequence 365 AA;

Query Match 58.4%; Score 1080.5; DB 16; Length 365;
Best Local Similarity 59.7%; Pred. No. 4.2e-105;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFILFVFTVSTIFFHVOQR-----LAKIQAMWELPQIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLHIHQDSFPHGLGSLILCPDRRLVTPPAIFCLPGTAMGNASSC 70
QY 54 ALGPSQLRGMTTINAGRLGNOMGEYATLALAKNGRPAPFAPOMHSTLAPIFRITLPV 113
DB 71 PQHPASLSGTWTVPNGRFGNMGQYATLLAQLNGRRRAFLPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPWNYHLNDWMEEEYRHPPGCVYVFTGPGCSWTFFYHHLRQELIOEFILHD 173
DB 131 LAPEVDSRTFWRELQLDHWMSEYADL-RDFFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREAAQKFLRGLQV--NGSRPGTFVGVHVRGDDYLVHMPKWKGVVADRYLQOALDWF 231
DB 190 HLREAAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVHMPKWKGVVADRYLQOALDWF 249
QY 232 RARYSSLIFFVTSNGMAWCARENIDTSHGDDVVFAGDGTGEGSPAKDFALLTCQNHITMIGT 291
DB 250 RAREAPVFWVTSNGMAWCARENIDTSGQDVTTFAGDGEATPWKDFALLTCQNHITMIGT 309
QY 292 FGIWAAYLTGGDTIYLANFTLPDSEFLKIKPKEAFLPEWGTGIAADLSPL 341
DB 310 FGFWAAYLAGSDTYVLANFTLPDSEFLKIKPKEAFLPEWGTGIAADLSPL 359

RESULT 14
AAR70421
ID AAR70421 standard; Protein; 365 AA.
XX AAR70421;
XX DT 27-FEB-1996 (first entry)
XX DE 2-Alpha-fucosyltransferase.
XX KW GDP-L-fucose-beta-D-galactoside 2-alpha-fucosyltransferase;
XX KW glycosyltransferase; oligosaccharide; glycoprotein; glycolipid;
XX KW transgenic animal; cattle; glycosylation; milk.

```

```

XX OS Homo sapiens.
XX PN W09524488-A1.
XX XX
XX PD 14-SEP-1995.
XX PF 24-JAN-1995; 95WO-US01147.
XX PR 09-MAR-1994; 94US-0209132.
XX PA (ABBO ) ABBOTT LAB.
XX PI Cummings RD, Kopchik JJ, Moremen KW, Mukerji P;
XX PI Pierce JM, Prieto PA, Smith DF;
XX DR WPI; 1995-328279/42.
XX DR N-PSDB; AAT01082.
XX PT Transgenic animal expressing heterologous catalyst - used in
XX PT metabolic prodn esp. glycosyltransferase for prodn. of
XX PT oligosaccharide(s) and glyco-conjugate(s) in the milk, useful in
XX PT foods, pharmaceuticals, etc.
XX PS Example 1; Page 62-64; 84pp; English.
XX CC Alpha-1,2-fucosyltransferase (AAR70421) is encoded by a cDNA clone
XX CC (AAT01082) isolated from a human epidermal carcinoma cell line A431
XX CC cDNA library. The enzyme can be expressed in the milk of a
XX CC transgenic mammal, esp. cow. This allows large-scale prodn. of
XX CC oligosaccharides or glycosylated proteins and lipids in the milk.
XX SQ Sequence 365 AA;

Query Match 58.4%; Score 1080.5; DB 16; Length 365;
Best Local Similarity 59.7%; Pred. No. 4.2e-105;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFILFVFTVSTIFFHVOQR-----LAKIQAMWELPQIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLHIHQDSFPHGLGSLILCPDRRLVTPPAIFCLPGTAMGNASSC 70
QY 54 ALGPSQLRGMTTINAGRLGNOMGEYATLALAKNGRPAPFAPOMHSTLAPIFRITLPV 113
DB 71 PQHPASLSGTWTVPNGRFGNMGQYATLLAQLNGRRRAFLPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPWNYHLNDWMEEEYRHPPGCVYVFTGPGCSWTFFYHHLRQELIOEFILHD 173
DB 131 LAPEVDSRTFWRELQLDHWMSEYADL-RDFFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREAAQKFLRGLQV--NGSRPGTFVGVHVRGDDYLVHMPKWKGVVADRYLQOALDWF 231
DB 190 HLREAAQSVLGQLRLGRTGDRPRTFVGVHVRGDDYLVHMPKWKGVVADRYLQOALDWF 249
QY 232 RARYSSLIFFVTSNGMAWCARENIDTSHGDDVVFAGDGTGEGSPAKDFALLTCQNHITMIGT 291
DB 250 RAREAPVFWVTSNGMAWCARENIDTSGQDVTTFAGDGEATPWKDFALLTCQNHITMIGT 309
QY 292 FGIWAAYLTGGDTIYLANFTLPDSEFLKIKPKEAFLPEWGTGIAADLSPL 341
DB 310 FGFWAAYLAGSDTYVLANFTLPDSEFLKIKPKEAFLPEWGTGIAADLSPL 359

RESULT 15
AAR90572
ID AAR90572 standard; Protein; 365 AA.
XX AAR90572;
XX DT 08-APR-1996 (first entry)
XX DE Human H-transferase.

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:03:22 : Search time 14.1436 Seconds
(without alignments)
715.625 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVQMPFSPFMAHFLVF.....AAFLPWTGIAADLSPLLK 344

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1850	100.0	344	4	US-09-298-886-11
2	1832.5	99.1	343	1	US-08-395-800A-8
3	1832.5	99.1	343	1	US-09-254-077A-8
4	1528	82.6	340	4	US-09-254-077A-6
5	1528	82.6	340	4	US-09-254-077A-7
6	1509.5	81.6	347	4	US-09-254-077A-9
7	1475	79.7	380	4	US-09-298-886-8
8	1376	74.4	353	4	US-09-298-886-10
9	1093.5	59.1	373	4	US-09-254-077A-12
10	1080.5	58.4	365	1	US-07-914-281-6
11	1080.5	58.4	365	1	US-08-393-246-6
12	1080.5	58.4	365	1	US-08-273-411-1
13	1080.5	58.4	365	1	US-08-525-058A-6
14	1080.5	58.4	365	1	US-08-395-800A-6
15	1080.5	58.4	365	1	US-08-395-800A-10
16	1080.5	58.4	365	2	US-08-696-731-6
17	1080.5	58.4	365	4	US-09-042-531-6
18	1080.5	58.4	365	5	PCT-US91-00899-11
19	1080.5	58.4	365	5	PCT-US91-00899-12
20	1076.5	58.2	333	5	PCT-US91-00899-11
21	1075.5	58.1	365	4	US-09-151-592-2
22	1075.5	58.1	365	4	US-09-254-077A-10
23	809.5	43.8	222	1	US-08-395-800A-2
24	420	22.7	102	1	US-08-395-800A-3
25	144.5	7.8	300	4	US-09-433-598-2
26	100	5.4	355	4	US-09-390-131-5
27	92	5.0	575	3	US-08-913-805A-2

28	92	5.0	575	4	US-09-442-629-2	Sequence 2, Appli
29	90	4.9	575	3	US-08-913-805A-10	Sequence 10, Appl
30	90	4.9	575	4	US-09-442-629-10	Sequence 10, Appl
31	87.5	4.7	365	3	US-08-978-741-2	Sequence 2, Appli
32	87.5	4.7	365	4	US-09-333-729A-3	Sequence 3, Appli
33	87.5	4.7	397	3	US-08-978-741-6	Sequence 6, Appli
34	87.5	4.7	397	4	US-09-333-729A-7	Sequence 7, Appli
35	82.5	4.5	183	1	US-08-167-035-33	Sequence 33, Appl
36	82.5	4.5	183	2	US-08-208-887A-33	Sequence 33, Appl
37	82.5	4.5	183	2	US-08-539-005-33	Sequence 33, Appl
38	82.5	4.5	183	4	US-09-280-598-35	Sequence 35, Appl
39	82.5	4.5	1048	4	US-08-887-534A-85	Sequence 85, Appl
40	81.5	4.4	607	1	US-07-959-943-7	Sequence 7, Appli
41	81.5	4.4	630	1	US-07-959-943-9	Sequence 9, Appli
42	81.5	4.4	653	1	US-07-782-298-2	Sequence 2, Appli
43	80	4.3	604	2	US-08-735-041A-4	Sequence 4, Appli
44	80	4.3	604	3	US-09-190-476B-4	Sequence 4, Appli
45	80	4.3	604	3	US-09-190-889A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-298-886-11
; Sequence 11, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-886-11

Query Match 100.0%; Score 1850; DB 4; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.9e-204;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLVQMPFSPFMAHFLVFVTYTIHFVQORLAKIQAMHQLPVQIPVLASTSKALGPSOL	60
DB	1	MLVQMPFSPFMAHFLVFVTYTIHFVQORLAKIQAMHQLPVQIPVLASTSKALGPSOL	60
QY	61	RGMTINAIGRGNOMGEYATLYALAKMNGRPAFIPAQMHSITLAPIFRITLPLVLSATAS	120
DB	61	RGMTINAIGRGNOMGEYATLYALAKMNGRPAFIPAQMHSITLAPIFRITLPLVLSATAS	120
QY	121	RIPWQNYHLNDMEYRIHPGEVYRTGYPCSWTFYHHLRQEIQLQETLHDHVRBAQ	180
DB	121	RIPWQNYHLNDMEYRIHPGEVYRTGYPCSWTFYHHLRQEIQLQETLHDHVRBAQ	180
QY	181	KFLRLQVNGSRPGFVGVVHVRGDYVHVMPKVGKVGVVADRYLQALDWFARVSSLIIF	240
DB	181	KFLRLQVNGSRPGFVGVVHVRGDYVHVMPKVGKVGVVADRYLQALDWFARVSSLIIF	240
QY	241	VVTSNGMAWCNENIDTSHGDVVVAGDGIEGSPAKDFALLITQCNHITMTIGTIGIWAAYLT	300
DB	241	VVTSNGMAWCNENIDTSHGDVVVAGDGIEGSPAKDFALLITQCNHITMTIGTIGIWAAYLT	300
QY	301	GGDTIYLANYLTPSPFLKIFKPEAAFLPWTGIAADLSPLLK 344	
DB	301	GGDTIYLANYLTPSPFLKIFKPEAAFLPWTGIAADLSPLLK 344	

RESULT 2

US-08-395-800A-8
 ; Sequence 8, Application US/08395800A
 ; Patent No. 5807732

GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B
 ; APPLICANT: LENNON, GREGORY
 ; APPLICANT: ROQUIER, SYLVIE
 ; APPLICANT: GIORGI, DOMINIQUE
 ; APPLICANT: KELLY, ROBERT J
 ; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
 ; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES/DNA SEQUENCES ENCODING THE
 ; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
 ; TITLE OF INVENTION: GENOTYPING A PERSON
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEFAX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 343 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-395-800A-8

Query Match 99.1%; Score 1832.5; DB 1; Length 343;
 Best Local Similarity 99.7%; Pred. No. 5e-202;
 Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60

Db 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60

QY 61 RGMWTINAGRIGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAPIFRITTLPVLSATAS 120

Db 61 RGMWTINAGRIGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAPIFRITTLPVLSATAS 120

QY 121 RIPWQNYHLNDWMEEYRHIIPGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 180

Db 121 RIPWQNYHLNDWMEEYRHIIPGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 179

QY 181 KFLRGLQVNGSRPGTFVGVHVRGDIYVHMPKWKGVVADRRYLOQALDWFRRARYSSLIF 240

Db 180 KFLRGLQVNGSRPGTFVGVHVRGDIYVHMPKWKGVVADRRYLOQALDWFRRARYSSLIF 239

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFGIWAAYLT 300

Db 240 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFGIWAAYLT 299

QY 301 GGDITYLANITYLPDSPFLKIFKPEAAFLPEWTGTIAADLSPLKH 344

Db 300 GGDITYLANITYLPDSPFLKIFKPEAAFLPEWTGTIAADLSPLKH 343

RESULT 3

US-09-254-077A-8

; Sequence 8, Application US/09254077A

; Patent No. 6399758

GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.
 ; APPLICANT: MCKENZIE, IAN C. F.
 ; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; FILE REFERENCE: 30562.5USWO
 ; CURRENT APPLICATION NUMBER: US/09/254.077A
 ; CURRENT FILING DATE: 1999-06-11
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00540
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: PO 1823
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-254-077A-8

Query Match 99.1%; Score 1832.5; DB 4; Length 343;
 Best Local Similarity 99.7%; Pred. No. 5e-202;
 Matches 343; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60

Db 1 MLVQMPFSFPMAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60

QY 61 RGMWTINAGRIGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAPIFRITTLPVLSATAS 120

Db 61 RGMWTINAGRIGNOMGEYATLYALAKMNGRPAFIPAQMHSHTLAPIFRITTLPVLSATAS 120

QY 121 RIPWQNYHLNDWMEEYRHIIPGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 180

Db 121 RIPWQNYHLNDWMEEYRHIIPGEYVRFYPCSWTFYHHLRQELQETLHDHVREBAQ 179

QY 181 KFLRGLQVNGSRPGTFVGVHVRGDIYVHMPKWKGVVADRRYLOQALDWFRRARYSSLIF 240

Db 180 KFLRGLQVNGSRPGTFVGVHVRGDIYVHMPKWKGVVADRRYLOQALDWFRRARYSSLIF 239

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFGIWAAYLT 300

Db 240 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNHNTIMTIGTFGIWAAYLT 299

QY 301 GGDITYLANITYLPDSPFLKIFKPEAAFLPEWTGTIAADLSPLKH 344

Db 300 GGDITYLANITYLPDSPFLKIFKPEAAFLPEWTGTIAADLSPLKH 343

RESULT 4

US-09-254-077A-6

; Sequence 6, Application US/09254077A

; Patent No. 6399758

GENERAL INFORMATION:

; APPLICANT: SANDRIN, MAURO S.
 ; APPLICANT: MCKENZIE, IAN C. F.
 ; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
 ; FILE REFERENCE: 30562.5USWO
 ; CURRENT APPLICATION NUMBER: US/09/254.077A
 ; CURRENT FILING DATE: 1999-06-11
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00540
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: PO 1823
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Sus scrofa

US-09-254-077A-6

```
Query Match      82.6%; Score 1528; DB 4; Length 340;
Best Local Similarity 82.0%; Pred. No. 5e-167;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVOIPVLASTSKALGPSOL 60
DB 1 MLSMQASFFFTGPGFLFVFTASTIFHLQORWVKIQPTWELQM--VTQVTTESPPQL 57

QY 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAFIPAQMNSTLAPIFRITLPLVLSATAS 120
DB 58 KGMWTINAIGRLGNQMGEXATLYALAKMNGRPAFIPPEMNSTLAPIFRITLPLVLSATAS 117

QY 121 RIPQWNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVEEAQ 180
DB 118 RIPQWNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVEEAQ 176

QY 181 KFLRGLQVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYSSLI 240
DB 177 DFLRGLRVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYSPVF 236

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCGIWA 300
DB 237 VVSSNGMAWCRENINASRGDVFAGNIEGSPAKDFALLTQCNTHTIMTIGTFCGIWA 296

QY 301 GGTIYLANVTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLK 344
DB 297 GGTIYLANVTLDPSPFLKIFKPEAAFLPEWIGIEADLSPLLK 340

RESULT 5
US-09-254-077A-7
; Sequence 7, Application US/09254077A
; Patent No. 6399758
; GENERAL INFORMATION:
; APPLICANT: SANDRIN, MAURO S.
; TITLE OF INVENTION: NUCLEIC ACIDS FOR REDUCING CARBOHYDRATE EPITOPES
; FILE REFERENCE: 30562.SUSWO
; CURRENT APPLICATION NUMBER: US/09/254,077A
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: PCT/AU97/00540
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: PO 1823
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-254-077A-7

Query Match      82.6%; Score 1528; DB 4; Length 340;
Best Local Similarity 82.0%; Pred. No. 5e-167;
Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVOIPVLASTSKALGPSOL 60
DB 1 MLSMQASFFFTGPGFLFVFTASTIFHLQORWVKIQPTWELQM--VTQVTTESPPQL 57

QY 61 RGMWTINAIGRLGNQMGEXATLYALAKMNGRPAFIPAQMNSTLAPIFRITLPLVLSATAS 120
DB 58 KGMWTINAIGRLGNQMGEXATLYALAKMNGRPAFIPPEMNSTLAPIFRITLPLVLSATAS 117

QY 121 RIPQWNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVEEAQ 180
DB 118 RIPQWNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVEEAQ 176

QY 181 KFLRGLQVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYSSLI 240
DB 177 DFLRGLRVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYSPVF 236

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCGIWA 300

Query Match      81.6%; Score 1509.5; DB 4; Length 347;
Best Local Similarity 81.3%; Pred. No. 6.9e-165;
Matches 283; Conservative 18; Mismatches 42; Indels 5; Gaps 3;

QY 1 MLVQMPFSPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVOIPVLASTSKALG 56
DB 1 MSTAQVFPFAPFVHVILFVFTASTIFHLQORLVRQIQTWEELPALTPAVTFRTSQRAP 60

QY 57 PSOLRGMTINAIGRLGNQMGEXATLYALAKMNGRPAFIPAQMNSTLAPIFRITLPLVLS 116
DB 61 SRPLGGMWTINAMGRGNQMGEXATLYALAKMNGRPAFIPAQMNSTLAPIFRISLPLVLS 120

QY 117 ATASRIPWQNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVR 176
DB 121 STASRVPWQNLNDWMEERYRHIPGCEYVFTGCPSCWTFYHHLRQEILOEFTLHDHVR 179

QY 177 EEAQKFLRGLQVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYS 236
DB 180 EEAQKFLRGLRVNSRGFTFVGVHVRGDDYVHVMKPVKGVADRRYLOQALDWFRRARYR 239

QY 237 SLIPVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCGIWA 296
DB 240 SPVFTVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFCGIWA 299

QY 297 AYLTTGGTIVIANVTLDPSPFLKIFKPEAAFLPEWTGIAADLSPLLK 344
DB 300 AYLTTGGTIVIANVTLDPSPFLKIFKPEAAFLPEWIGINADLSPLLK 347

RESULT 7
US-09-298-886-8
; Sequence 8, Application US/09298886
; Patent No. 6329170
; GENERAL INFORMATION:
; APPLICANT: Eric H. Holmes et al.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
; TITLE OF INVENTION: GM1-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 8511-029
; CURRENT APPLICATION NUMBER: US/09/298,886
; CURRENT FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 29
```


;; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,
;; ADDRESSEE: P.C.
;; STREET: 1755 Jefferson Davis Highway, Fourth Floor
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/914,281
;; FILING DATE: 19920720
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lavalleye, Jean-Paul M. P.
;; REGISTRATION NUMBER: 31,451
;; REFERENCE/DOCKET NUMBER: 2363-060-55
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)521-4500
;; TELEFAX: (703)486-2347
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 365 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
US-07-914-281-5

Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;
QY 12 MAHFLFVETVSTIFVHOQR-----LAKIQAMWELPVOIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLIHODSPHGLSILCPDRLVTPPVAIFCLPGTAMGNASSSC 70
QY 54 ALGPSOLRGMTINAIIGRLGNOMGEYATLYALAKMNGRPAFIPAQMNSTLAPFRITLPV 113
DB 71 PQHPASLSGTWVTPNGRFGNQMGQVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPQWNYHLNDKMEYRHPGGEYVRFYGYPCSTFYHHLRQEILOEFTLHD 173
DB 131 LAPEVDSRTPWRELQJLHDWSEYADL-RDPFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREEAQKFLRGLOV--NGSRPGTFVGVHVRGDIYVMPKVKGVVADRYLQQAQDWF 231
DB 190 HLREEAQSVLGQLRGTRPRTFVGVHVRGDIYVMPKVKGVVADRYLQQAQDWF 249
QY 232 RARYSSLIFVVTNSGMWCRENIDTSGDVTFAAGDIEGSPAKDFALLTQCNTIMTIGT 291
DB 250 RARHEAPVFTVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNTIMTIGT 309
QY 292 FGIWAAYLTGGDIYIYLANITLPDSPFLKIFKPEAAFLPEWTGIAADLSPL 341
DB 310 FGFWAAYLAGGDIYIYLANITLPDSPFLKIFKPEAAFLPEWVGINADLSPL 359

RESULT 11
US-08-393-246-6
; Sequence 6, Application US/08393246
; Patent No. 559590
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; AND FOR THE ISOLATION
; OF GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

;; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,
;; ADDRESSEE: P.C.
;; STREET: 1755 Jefferson Davis Highway, Fourth Floor
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/393,246
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/220,433
;; FILING DATE: 30-MAR-1994
;; APPLICATION NUMBER: US 07/914,281
;; FILING DATE: 20-JUL-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lavalleye, Jean-Paul M. P.
;; REGISTRATION NUMBER: 31,451
;; REFERENCE/DOCKET NUMBER: 2363-060-55
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)521-4500
;; TELEFAX: (703)486-2347
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 365 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
US-08-393-246-6

Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;
QY 12 MAHFLFVETVSTIFVHOQR-----LAKIQAMWELPVOIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLIHODSPHGLSILCPDRLVTPPVAIFCLPGTAMGNASSSC 70
QY 54 ALGPSOLRGMTINAIIGRLGNOMGEYATLYALAKMNGRPAFIPAQMNSTLAPFRITLPV 113
DB 71 PQHPASLSGTWVTPNGRFGNQMGQVATLLAQLNGRRAFILPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPQWNYHLNDKMEYRHPGGEYVRFYGYPCSTFYHHLRQEILOEFTLHD 173
DB 131 LAPEVDSRTPWRELQJLHDWSEYADL-RDPFLKLSGFPSCSWTFHHLRQIRREFTLHD 189
QY 174 HVREEAQKFLRGLOV--NGSRPGTFVGVHVRGDIYVMPKVKGVVADRYLQQAQDWF 231
DB 190 HLREEAQSVLGQLRGTRPRTFVGVHVRGDIYVMPKVKGVVADRYLQQAQDWF 249
QY 232 RARYSSLIFVVTNSGMWCRENIDTSGDVTFAAGDIEGSPAKDFALLTQCNTIMTIGT 291
DB 250 RARHEAPVFTVTSNGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQCNTIMTIGT 309
QY 292 FGIWAAYLTGGDIYIYLANITLPDSPFLKIFKPEAAFLPEWTGIAADLSPL 341
DB 310 FGFWAAYLAGGDIYIYLANITLPDSPFLKIFKPEAAFLPEWVGINADLSPL 359

RESULT 12
US-08-273-411-1
; Sequence 1, Application US/08273411
; Patent No. 5625124

Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 1.5e-115;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY	12	MAHFILFVETSTIFHVQOR-----LAKIQAMWELPVOIPVL-----ASTSK	53
Db	11	LAFLVLCVSVIFFLHIHQDSPHGLGLSILCPDRRLVTPPAVTCFLPGTAMGNASSC	70
QY	54	ALGPSOLRGMTTINAIGRLGNOMGEVATLYALAKNNGRPAPIPQAMHSTLAPIRITLPV	113
Db	71	POHPASLSGTIVYNGREGNOMGOYATLLAOLNGRRARFILPAMHAALAPVPRITLPV	130
QY	114	LHSATASRIPQNYHLNDWMEEEVYHIPPGEYVRFYGYPCSWTFYVHHLRQBILOEFTLHD	173
Db	131	LAPEVDSRTPMRELQLDHDMSEYADL--RDPEFLKSLGFPCCSWTFPHHLRQIRREFTLHD	189
QY	174	HYREAPQAFRLGLOW--NGSRPGTFVGVHVRGQYVYVMPKVKGVGVADRRYLOQALDNF	231
Db	190	HLREAPQSVLQALRGTRGDRPTFVGVHVRGDIQLQVPMQWAGVVGDSYAIRQAMDWF	249
QY	232	RARYSLLFVYTSNGMAWCRNITSDTSGDVVVFAGDIEGSPAKDFALLTCQCNITMTIGT	291

Db 250 RAREAPVFTVTSNGMEWCKENIDTSQDVTTFAGDQGEATPWKDFALLTCQNHMTIMTIGT 309

Qy 292 FGIWAAYLTGDDTYLYANLTLPDSPFLKIFKPEAAFLPEWGTGIADLSPL 341

Db 310 FGWAAYLAGDGTIVLANFTLPDSEFLKIFKPEAAFLPEWGTGINADLSPL 359

RESULT 14

US-08-395-800A-6

; Sequence 6, Application US/08395800A

; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; TITLE OF INVENTION: GENOTYPING A PERSON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 365 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-395-800A-6

Query Match 58.4%; Score 1080.5; DB 1; Length 365;

Best Local Similarity 59.7%; Pred. No. 1.5e-115;

Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

Qy 12 MAHFILFVFTVSTIFVHOOR-----LAKIOAMWELPVQIPVL-----ASTSK 53

Db 11 LAFLLVCVLSVIFFLHQQDSFPHGLGLSLCPDRRLVTPPVAIFCLPGTAMGNASSC 70

Qy 54 ALGPSQLRGWMTNATIGRLGNQMGYATLYALAKMNGRPAPFIPAQMSTLAPIFRITLPV 113

Db 71 PQHPASLSGTWTVYPNGRFGNQGYATLLAQLNGRRAPFILPAMHAALAPVFRITLPV 130

Qy 114 LHSATASRIPQNYHLNDWMEEEYRHPPGEYVFTGYPGCSWTFYHHLRQEILOEFTLHD 173

Db 131 LAPVDSTPTWRELQHLDMWSEYADL-RDPFLKLSGFPSCSWTFHHLRQEIREFTLHD 189

Qy 174 HVREAAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVMKVKWGVVADRYLQOALDNF 231

Db 190 HLREAAQSVLQGLRLGTGDRPRTFVGVHVRGDDYLVMPQRMKGVVGDAYSAYLRQAMDWF 249

Qy 232 RARYSSLIFFVVTSGMWCARENIDTSHGDVVFAGDGIEGSPAKDFALLTCQNHMTIMTIGT 291

Db 250 RAREAPVFTVTSNGMEWCKENIDTSQDVTTFAGDQGEATPWKDFALLTCQNHMTIMTIGT 309

Qy 292 FGIWAAYLTGDDTYLYANLTLPDSPFLKIFKPEAAFLPEWGTGIADLSPL 341

Db 310 FGWAAYLAGDGTIVLANFTLPDSEFLKIFKPEAAFLPEWGTGINADLSPL 359

RESULT 15

US-08-395-800A-10

; Sequence 10, Application US/08395800A

; Patent No. 5807732

; GENERAL INFORMATION:

; APPLICANT: LOWE, JOHN B

; APPLICANT: LENNON, GREGORY

; APPLICANT: ROQUIER, SYLVIE

; APPLICANT: GIORGI, DOMINIQUE

; APPLICANT: KELLY, ROBERT J

; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE

; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE

; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF

; TITLE OF INVENTION: GENOTYPING A PERSON

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/395,800A

; FILING DATE: 28-FEB-1995

; CLASSIFICATION: 435

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 365 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-395-800A-10

Query Match 56.4%; Score 1080.5; DB 1; Length 365;

Best Local Similarity 59.7%; Pred. No. 1.5e-115;

Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

Qy 12 MAHFILFVFTVSTIFVHOOR-----LAKIOAMWELPVQIPVL-----ASTSK 53

Db 11 LAFLLVCVLSVIFFLHQQDSFPHGLGLSLCPDRRLVTPPVAIFCLPGTAMGNASSC 70

Qy 54 ALGPSQLRGWMTNATIGRLGNQMGYATLYALAKMNGRPAPFIPAQMSTLAPIFRITLPV 113

Db 71 PQHPASLSGTWTVYPNGRFGNQGYATLLAQLNGRRAPFILPAMHAALAPVFRITLPV 130

Qy 114 LHSATASRIPQNYHLNDWMEEEYRHPPGEYVFTGYPGCSWTFYHHLRQEILOEFTLHD 173

Db 131 LAPVDSTPTWRELQHLDMWSEYADL-RDPFLKLSGFPSCSWTFHHLRQEIREFTLHD 189

Qy 174 HVREAAKFLRGLOV--NGSRPGTFVGVHVRGDDYVHVMKVKWGVVADRYLQOALDNF 231

Db 190 HLREAAQSVLQGLRLGTGDRPRTFVGVHVRGDDYLVMPQRMKGVVGDAYSAYLRQAMDWF 249

Qy 232 RARYSSLIFFVVTSGMWCARENIDTSHGDVVFAGDGIEGSPAKDFALLTCQNHMTIMTIGT 291

Db 250 RAREAPVFTVTSNGMEWCKENIDTSQDVTTFAGDQGEATPWKDFALLTCQNHMTIMTIGT 309

QY 292 FGINAAYLTGGDTIYLANFTLPDSEFLKIFKPEAAFLPEWTGIAADLSPL 341
DB 310 FGFWAAYLAGDVTYLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL 359

Search completed: May 27, 2003, 15:11:53
Job time : 15.1436 secs

Result No.	Query			ID	Description
	Score	Match	Length		
1	1850	100.0	344	10	US-09-999-672-11
2	1850	100.0	344	12	US-10-040-863-11
3	1588	82.6	340	10	US-09-051-034A-2
4	1475	79.7	380	10	US-09-999-672-8
5	1475	79.7	380	12	US-10-040-863-8
6	1376	74.4	353	10	US-09-999-672-10
7	1376	74.4	353	12	US-10-040-863-10
8	1081.5	58.5	365	10	US-09-844-268-13
9	1081.5	58.5	365	10	US-09-844-705-13
10	1080.5	58.4	365	9	US-10-105-563-10
11	1080.5	58.4	365	10	US-09-863-475A-6
12	1075.5	58.1	365	10	US-09-051-034A-4
13	144.5	7.8	300	10	US-09-848-838-2
14	119.5	6.5	431	9	US-09-479-614-14
15	119.5	6.5	496	9	US-09-479-614-2
16	119.5	6.5	496	9	US-09-479-614-29
17	119.5	6.5	496	9	US-10-214-524-25
18	92	5.0	575	10	US-09-839-136-2
19	91	4.9	575	10	US-09-839-136-10

QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300
 |||||
 Db 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300
 |||||
 QY 301 GGDITVLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 344
 |||||
 Db 301 GGDITVLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 344
 |||||

RESULT 2
 US-10-040-863-11
 ; Sequence 11, Application US/10040863
 ; Patent No. US20020137165A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eric H. Holmes et al.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
 ; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
 ; FILE REFERENCE: 8511-029
 ; CURRENT APPLICATION NUMBER: US/10/040,863
 ; PRIOR FILING DATE: 2001-11-01
 ; PRIOR APPLICATION NUMBER: 09/298,886
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 11
 ; LENGTH: 344
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-040-863-11

Query Match 100.0%; Score 1850; DB 12; Length 344;
 Best Local Similarity 100.0%; Pred. No. 9.1e-183;
 Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
 |||||
 Db 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
 |||||
 QY 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHATAS 120
 |||||
 Db 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHATAS 120
 |||||
 QY 121 RIPWNYHLNDWMEERYRHIPPGYVRFYVPCSWTFYHHLRQELQEFTHDHRVREAAQ 180
 |||||
 Db 121 RIPWNYHLNDWMEERYRHIPPGYVRFYVPCSWTFYHHLRQELQEFTHDHRVREAAQ 180
 |||||
 QY 181 KFLRGLOVNSRPGTFVGVHVRGDYVHVMPKVKGVVADRRYLOQALDWFRRYSSLIIF 240
 |||||
 Db 181 KFLRGLOVNSRPGTFVGVHVRGDYVHVMPKVKGVVADRRYLOQALDWFRRYSSLIIF 240
 |||||
 QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300
 |||||
 Db 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300
 |||||
 QY 301 GGDITVLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 344
 |||||
 Db 301 GGDITVLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 344
 |||||

RESULT 3
 US-09-051-034A-2
 ; Sequence 2, Application US/09051034A
 ; Patent No. US20010055584A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL
 ; APPLICANT: SANDRIN, MAURO SERGIO
 ; TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC
 ; FILE REFERENCE: 30562.6USWO
 ; CURRENT APPLICATION NUMBER: US/09/051,034A
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: PCT/AU97/00492

; PRIOR FILING DATE: 1997-08-01
 ; PRIOR APPLICATION NUMBER: 60/024,279
 ; PRIOR FILING DATE: 1996-08-21
 ; PRIOR APPLICATION NUMBER: P01402
 ; PRIOR FILING DATE: 1996-08-02
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 340
 ; TYPE: PRT
 ; ORGANISM: Sus Domesticus
 US-09-051-034A-2

Query Match 82.6%; Score 1528; DB 10; Length 340;
 Best Local Similarity 82.0%; Pred. No. 1.7e-149;
 Matches 282; Conservative 24; Mismatches 34; Indels 4; Gaps 2;

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
 |||||
 Db 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
 |||||
 QY 61 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHATAS 120
 |||||
 Db 58 RGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHATAS 117
 |||||
 QY 121 RIPWNYHLNDWMEERYRHIPPGYVRFYVPCSWTFYHHLRQELQEFTHDHRVREAAQ 180
 |||||
 Db 118 RIPWNYHLNDWMEERYRHIPPGYVRFYVPCSWTFYHHLRQELQEFTHDHRVREAAQ 176
 |||||
 QY 181 KFLRGLOVNSRPGTFVGVHVRGDYVHVMPKVKGVVADRRYLOQALDWFRRYSSLIIF 240
 |||||
 Db 177 DFLKGLRVNGSRPSTYGVHVRGDYVHVMPKVKGVVADRRYLOQALDWFRRYSSLIIF 236
 |||||
 QY 241 VVTSNGMAWCRENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 300
 |||||
 Db 237 VVSSNGMAWCRENINASRGDDVVEAGNGIEGSPAKDFALLTQCNTIMTIGTFGIWAAYLT 296
 |||||
 QY 301 GGDITVLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 344
 |||||
 Db 297 GGETIYLANITLSDPSFLKIFKPEAAFLPEWTGIAADLSPLKX 340
 |||||

RESULT 4
 US-09-999-672-8
 ; Sequence 8, Application US/09999672
 ; Patent No. US20020127655A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eric H. Holmes et al.
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF A RAT GANGLIOSIDE
 ; TITLE OF INVENTION: GMI-SPECIFIC ALPHA1-2 FUCOSYLTRANSFERASE AND USES
 ; FILE REFERENCE: 8511-029
 ; CURRENT APPLICATION NUMBER: US/09/999,672
 ; PRIOR FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: US/09/298,886
 ; PRIOR FILING DATE: 1999-04-26
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-999-672-8

Query Match 79.7%; Score 1475; DB 10; Length 380;
 Best Local Similarity 77.7%; Pred. No. 5.9e-144;
 Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPFSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSQL 60
 |||||
 Db 1 MASAOVPFSPFLAHLFVFTVSTIHLQORIVKIQPLSEKELPWTQMSSTNTSPEHR 60
 |||||
 QY 50 STSKALGPSQLRGMWTINAGRLGNQMGAYATLYALAKMNGRPAFIPQMHSTLAPIFRITLPLVLSHATAS 109
 |||||


```

QY 332 TGIADLSPL 341
      || |||||
Db 350 VGINADLSPL 359

RESULT 9
US-09-844-705-13
; Sequence 13, Application US/09844705
; Patent No. US2002013936A1
; GENERAL INFORMATION:
; APPLICANT: BOSWORTH, BRAD
; APPLICANT: VOGELI, PETER
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
; TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
; FILE REFERENCE: 21419/90368
; CURRENT APPLICATION NUMBER: US/09/844,705
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 09/443,766
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Porcine
US-09-844-705-13

```

Query Match	58.5%;	Score	1081.5;	DB	10;	Length	365;
Best Local Similarity	65.5%;	Pred. No.	2.6e-103;				
Matches	203;	Conservative	30;	Mismatches	66;	Indels	11;
QY	42	PVOIPVLA	ST-----SKALGFSOLRGWNTNATGRGNQMGVEYATYIAJAKMGRPA	93			
Db	51	PVAIFCLAGTVPV	PNASDCSPKHPASFGTWTIYDPPGRFGNQMGVATYALLAQLNGRQA	110			
QY	94	FIAQKHSIIAP	FRITLPIVLHSNTASRIPQNYHLNDHMEERYHIPPGEYVRFYGYPC	153			
Db	111	FIOPNMHVLAP	FRITLPIVLAEVDHAPWRELEHDMWSEYAHKBPWIKLGFPC	169			
QY	154	SWTFYHHLRGEI	LQEEFTLHDHVEEAQKFLRGQV--NGSRPQTFVGVHVRGDDYVHVMP	211			
Db	170	SWTFFHHLREQ	IRSEFTLHDHLHQAGGVLSQFRLPRTGDRPSTFGVGHVRGDIYLRVP	229			
QY	212	KWKGVVADRY	LQOALDWPFRYSSLIFFVTSNGMAWORENIDTSGHDVVRGADIGES	271			
Db	230	KRWKGVVDG	RYLQQAAMDVFRARYEAPFVYVTSNGMEWCNKINTDSRGDVIYFAGDGREA	289			
QY	272	PAKDFALLTC	NCNTIMTIGTFGWAAYLTGGDPIYLANVTLDPSPFLKFKPEAAPLPEW	331			
Db	290	PAKDFALLV	OCNTIMTIGTFGWAAYLAGGDIYLANVTLPTSSTFLKFKPEAAPLPEW	349			
QY	332	TGIAADLS	SPL	341			
Db	350	VGINADLS	SPL	359			

```

Db      290  PARDFALLVQCNTIMTIGTFGWAAYLAGGDDIYLANFTLTSSFLKFKPFAFLPEM 349
QY      332  TGIADLSPL 341
        || |||||
Db      350  VGINADLSPL 359

RESULT 10
US-10-105-963-10
; Sequence 10, Application US/10105963
; Publication No. US20030068818A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Dending, Chris
; APPLICANT: Clark, A. John
; APPLICANT: Schiff, J. Michael
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Select
; TITLE OF INVENTION: Recombination
; FILE REFERENCE: 731/002
; CURRENT APPLICATION NUMBER: US/10/105,963
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/277,811
; PRIOR FILING DATE: 2001-03-21

```

; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-10

Query Match 58.4%; Score 1080.5; DB 9; Length 365;
Best Local Similarity 59.7%; Pred. No. 3.3e-103;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFLVFTVSTIFHVQOR-----LAKIQAMWELPVQIPVL-----ASTSK 53
DB 11 LAFLVLCVLSVIFLHIHDSFPHGLSLICPDRLRVTPEVAIFCLPCTAMGNASSSC 70
QY 54 ALGPSQLRGMTINAIAGRLGNOMGEYATLYALAKNGRPAFIPAQMHSSTLAFIFRITLPV 113
DB 71 POHPASLSGTWTVYPNGRFGNQMGQYATLLAQLNGRAFLPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPQWYHNDWMEERYHIPPGEYVVRTGYPGCSWTFYHHLRQELQFTLHD 173
DB 131 LAPEVDSRTPRELQHLHDMSEYADL-RDPFLKLSGFPSCWTFYHHLRQELQFTLHD 189
QY 174 HVREAQKFLAGLQV--NGSRPGTFVGVHVRGDDVHVMPKVMKGVVADRRYLOQALDWF 231
DB 190 HLREAAQSVLGRLGRTRDPRTFVGVHVRGDDVHVMPKVMKGVVADRRYLOQALDWF 249
QY 232 RARYSSLFVVTNSGMACRENIDTSHGDDVVFAGDIGSPAKDFALLTQCNHTTMTIGT 291
DB 250 RARHEAPVFTVTSNGMECKENIDTSQGDVTFAGDQGEATPKDKFALLTQCNHTTMTIGT 309
QY 292 FGIWAAYLTGGDTIYLANFTLPDSFLKIFKPEAAFLPEWTCIAADLSPL 341
DB 310 FGFWAAYLAGGDTVYLANFTLPDSEFLKIFKPEAAFLPEWTCIAADLSPL 359

RESULT 11

US-09-863-475A-6
; Sequence 6, Application US/09863475A
; Patent No. US2002010268A1
; GENERAL INFORMATION:

APPLICANT: LOWE, JOHN B.

TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS

OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/863,475A

FILING DATE: 24-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/914,281

FILING DATE: 20-JUL-1992

ATTORNEY/AGENT INFORMATION:

NAME: Lavalleye, Jean-Paul M. P.

REGISTRATION NUMBER: 31,451

REFERENCE/DOCKET NUMBER: 2363-060-55

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-863-475A-6

Query Match 58.4%; Score 1080.5; DB 10; Length 365;
Best Local Similarity 59.7%; Pred. No. 3.3e-103;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;

QY 12 MAHFLVFTVSTIFHVQOR-----LAKIQAMWELPVQIPVL-----ASTSK 53
DB 11 LAFLVLCVLSVIFLHIHDSFPHGLSLICPDRLRVTPEVAIFCLPCTAMGNASSSC 70
QY 54 ALGPSQLRGMTINAIAGRLGNOMGEYATLYALAKNGRPAFIPAQMHSSTLAFIFRITLPV 113
DB 71 POHPASLSGTWTVYPNGRFGNQMGQYATLLAQLNGRAFLPAMHAALAPVFRITLPV 130
QY 114 LHSATASRIPQWYHNDWMEERYHIPPGEYVVRTGYPGCSWTFYHHLRQELQFTLHD 173
DB 131 LAPEVDSRTPRELQHLHDMSEYADL-RDPFLKLSGFPSCWTFYHHLRQELQFTLHD 189
QY 174 HVREAQKFLAGLQV--NGSRPGTFVGVHVRGDDVHVMPKVMKGVVADRRYLOQALDWF 231
DB 190 HLREAAQSVLGRLGRTRDPRTFVGVHVRGDDVHVMPKVMKGVVADRRYLOQALDWF 249
QY 232 RARYSSLFVVTNSGMACRENIDTSHGDDVVFAGDIGSPAKDFALLTQCNHTTMTIGT 291
DB 250 RARHEAPVFTVTSNGMECKENIDTSQGDVTFAGDQGEATPKDKFALLTQCNHTTMTIGT 309
QY 292 FGIWAAYLTGGDTIYLANFTLPDSFLKIFKPEAAFLPEWTCIAADLSPL 341
DB 310 FGFWAAYLAGGDTVYLANFTLPDSEFLKIFKPEAAFLPEWTCIAADLSPL 359

RESULT 12

US-09-051-034A-4

; Sequence 4, Application US/09051034A

; Patent No. US2001005584A1

; GENERAL INFORMATION:

APPLICANT: MCKENZIE, IAN FARQUHAR CAMPBELL

APPLICANT: SANDRIN, MAURO SERGIO

TITLE OF INVENTION: IMPROVED NUCLEIC ACIDS ENCODING A CHIMERIC

TITLE OF INVENTION: GLYCOSYLTRANSFERASE

FILE REFERENCE: 30562.6USWO

CURRENT APPLICATION NUMBER: US/09/051,034A

CURRENT FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: PCT/AU97/00492

PRIOR FILING DATE: 1997-08-01

PRIOR APPLICATION NUMBER: 60/024,279

PRIOR FILING DATE: 1996-08-21

PRIOR APPLICATION NUMBER: P01402

PRIOR FILING DATE: 1996-08-02

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 365

TYPE: PRT

ORGANISM: Sus Domesticus

US-09-051-034A-4

Query Match 58.1%; Score 1075.5; DB 10; Length 365;

Best Local Similarity 65.2%; Pred. No. 1.1e-102;

Matches 202; Conservative 30; Mismatches 67; Indels 11; Gaps 3;

QY 42 PQQIPVLAST-----SKALGPSQLRGMTINAIAGRLGNOMGEYATLYALAKNGRPA 93


```
Db 51 PVAIFCLAGTVPHPNADSCPKHFPASFGTWTIYDGRFGNMGQYATLLALALNGRQA 110
Qy 94 FIPAQHSTLAPIRITLPLVLHSATASRIPQNYHLNDWMEERYRHIPPGEYVRFYGPC 153
Db 111 FLOPMHVLAPVFRITLPLVLAPEDVDRHAPRELELHDMWSEDAHLAKE-PWLKLTGPPC 169
Qy 154 SWTFVHLRQEILOEFTLHDLHVREAAQKFLRGLOV--NGSRPGTFVGVHVRGDIYVHYMP 211
Db 170 SWTFPHLRQEIIRSEFTLHDLRQEAQGVLSQFRLPRIGDRPSIFVGVHVRGDIYVHYMP 229
Qy 212 KWKGVVADRYLQALDWFARVSSLIFFVTSNGMAWCRNIDTSHGDVVFAGDIGBS 271
Db 230 KWKGVGVDGAYLQAOAMDFAFYAPVFFVTSNGMCWCRNIDTSGDVIYFAGDGREAA 289
Qy 272 PAKDFALLTQCNTHTMTTGTGFWAAAYLTGDDTIVLANTPLDPSFPLKIPKPEAAFLPEW 331
Db 290 PARDFALLVQCNTHTMTTGTGFWAAAYLAGDGIYLANFTLPTSFLKIFRPEAAFLPEW 349
Qy 332 TGIADLSPL 341
Db 350 VGINADLSPL 359

RESULT 13
US-09-848-838-2
; Sequence 2, Application US/09848838
; Patent No. US20020037570A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-848-838-2

Query Match 7.8%; Score 144.5; DB 10; Length 300;
Best Local Similarity 22.9%; Pred. No. 1.3e-06;
Matches 61; Conservative 37; Mismatches 123; Indels 45; Gaps 8;

Qy 66 INAIGRLGNQMGEXATLYALAKMNGRPAPFA-----QMHSTLAPIRITLPLVLHSATAS 120
Db 6 VQICGGLGNQMGEXAFKSLQKHLNTPVLLDTTSEFDSNRKNQWLELFDIDLPYANAKA 65
Qy 121 RIPQNYHLNDWMEERYRHI-----PPGEYVRFYGPCSWTFYHLR 162
Db 66 IAKMQ--HLPKLVRLDALXIIGIDRVSRQEIYEPKILKPSRLTYFFCYQDPDYFDAIS 123
Qy 163 QEILQEEFTL-----HDHVREAAQKFLRGLOVNGSRPGTFVGVHVRGDIYVHYMPKV 213
Db 124 SLIKGTFTLPPPPNNKNNKKEEYQKLS--LTLAAKNVYF--VHIRRGDIYVGI----- 175
Qy 214 WKGVVADRYLQALDWFARVSSLIFFVTSNGMAWCRNIDTSHGDVVFAGDIGEGSPA 273
Db 176 --GCOLGIDYQKALEYNAKRVPMLEVFCEDIKAF-TQNLDLGYPTDMTRKEEBAY 232
Qy 274 KDFALLTQCNTHTMTTGTGFWAAAYL 299
Db 233 WMLLMQSKHGIIANSTYSWAAAYL 258

RESULT 14
Query Match 6.5%; Score 119.5; DB 9; Length 496;
Best Local Similarity 23.8%; Pred. No. 0.001;
Matches 72; Conservative 27; Mismatches 111; Indels 93; Gaps 15;
```

```
US-09-479-614-14
; Sequence 14, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-14

Query Match 6.5%; Score 119.5; DB 9; Length 431;
Best Local Similarity 23.8%; Pred. No. 0.00084;
Matches 72; Conservative 27; Mismatches 111; Indels 93; Gaps 15;

Qy 47 VLASTSKALGPSQLRGMWTINAIGRLGNQMGEXATLYALAKMNGRPAPFAIPQMHST---- 102
Db 102 VSACTMNFIPPTVKLFHSSCNPLGDTG-----STIQLCLISG---YVPGDMETWLVLD 152
Qy 103 ---LAPIFRITLPLVLHSATASRIPQNYHLNDWMEERYRHIPPGEYVRFYGPCSWTFYH 159
Db 153 GQKATNIFPYTAP--GKQEGKVTSTHSELN-----ITQGEWVSQKTYTCQVTY-- 198
Qy 160 HLRQEILOEFTLHDLHVREAAQKFLRGLOVNGSRPGTFVGVHVRGDIYVHYMPKVWKG 219
Db 199 -----QGTFEDHARKCTESDPRGVSTYLSPPSL-----DLIYHKSPKI-TCLV 243
Qy 220 DRYLQALDWFARVSSLIFFVTSNGM--AWCRNIDTSHGDVVFAGDIGEGSPAKDFA 277
Db 244 D-----LANTDGMILTSRENGESVHPD-----PMV 269
Qy 278 LITQCNHITMTIGTGIWAAYLTGDDTIVLANTPLDPS-----FLKI---FKPEA-A 326
Db 270 KKTQYNGTITVTSTLPVDATDWEGET-YQCKVTHPLPKDIVRSIAKAPGRPPPEVYV 328
Qy 327 FLP 329
Db 329 FLP 331

RESULT 15
US-09-479-614-2
; Sequence 2, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-2

Query Match 6.5%; Score 119.5; DB 9; Length 496;
Best Local Similarity 23.8%; Pred. No. 0.001;
Matches 72; Conservative 27; Mismatches 111; Indels 93; Gaps 15;
```

Wed May 28 09:22:07 2003

47	VLASTSKALGPSOLRGWMTTINAIGRLGNOMGEYATLYALAKMGRPAIPQAOMHST----	102
QY		
167	VSACIMNFIPPIVKLFHSSCNELGDTG-----STIQLLLISG---YVPDMEVILVWD	217
Db		
103	---LAPIFRITLPIVLHSATASRIPQNYHLNDWMEEYRHIPPGEYVFRFGPCSWTFYH	159
QY		
218	GOKATNIFYTAP---GKQEGKVTSTHSELN-----ITQGEWSVKYTYTCQVTV--	263
Db		
160	HLRQELQFTLHDHVREBAQKFLGVLQVNGSRPGTFVGVHVHRGDYHVMVKWKGYYVA	219
QY		
264	-----QGFTFEDHARKCTESDPCGVSTYLPSPPL-----DLYHKSEKI-TCLVW	308
Db		
220	DRYLOQALDWFRARYSSLIFFVVTNGM--AMCRENIDTSHGDVVFAGDIGSPAKDFA	277
QY		
309	D-----LANTDGMILTWSRENGESVHPD-----PMV	334
Db		
278	LLFQCNRHTIMTGTGFIWAAYTGGDTPIVLYANTLPDSP-----FLKI----FKPEA-A	326
QY		
335	KKTOYNGTINTVSTLPVDATDWMVEGT-YOCKVTHPDLPKDIVRSIAKAPGRFRFPPEVV	393
Db		
327	FLP	329
QY		
394	FLP	396
Db		

Search completed: May 27, 2003, 15:25:45
Job time : 16.8627 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:00:45 ; Search time 12.4655 Seconds
(without alignments)
2652.940 Million cell updates/sec

Title: US-10-040-863-11
Perfect score: 1850
Sequence: 1 MLVVQMPFSPMAHFILEVF.....AAFLPEWTGIAADLSPLLK 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1832.5	99.1	343	2 A56098	alpha(1,2)fucosylt
2	1375.5	74.4	354	2 B56392	beta-galactoside a
3	1093.5	59.1	373	2 A56392	beta-galactoside a
4	1080.5	58.4	365	2 A36047	galactoside 2-alpha
5	623	33.7	159	2 S46424	galactoside 2-alpha
6	593	32.1	142	2 S51582	galactoside 2-alpha
7	195	10.5	281	2 T44328	hypothetical prote
8	181.5	9.8	363	2 T20745	hypothetical prote
9	176	9.5	443	2 H87911	protein B0205.4 [i
10	155	8.4	395	2 T32309	hypothetical prote
11	144.5	7.8	348	2 T31916	hypothetical prote
12	141.5	7.6	335	2 T32294	hypothetical prote
13	138	7.5	299	2 H71976	probable alpha(1,2
14	137.5	7.4	392	2 T25334	hypothetical prote
15	131.5	7.1	348	2 T02798	hypothetical prote
16	131	7.1	625	2 T31051	hypothetical prote
17	129.5	7.0	388	2 T32307	hypothetical prote
18	122	6.6	365	2 T32553	hypothetical prote
19	120.5	6.5	383	2 T20572	hypothetical prote
20	120	6.5	434	2 T26275	hypothetical prote
21	118.5	6.4	353	2 T25390	hypothetical prote
22	118	6.4	500	2 T32068	hypothetical prote
23	115.5	6.2	277	2 T25307	hypothetical prote
24	115.5	6.2	365	2 T25309	hypothetical prote
25	114	6.2	381	2 T15140	hypothetical prote
26	101	5.5	1123	2 AD2045	hypothetical prote
27	100	5.4	355	2 T34405	hypothetical prote
28	93	5.0	371	2 T32692	hypothetical prote
29	91	4.9	317	2 T33887	hypothetical prote

30 91 4.9 339 2 T34406 hypothetical prote
31 91 4.9 401 2 B84828 hypothetical prote
32 90 4.9 575 1 JC5432 glycoprotein 6-alp
33 89.5 4.8 1462 1 S32437 pol polyprotein -
34 89 4.8 451 2 E72398 hypothetical prote
35 89 4.8 460 2 JC5137 beta-glucosidase (
36 89 4.8 1489 2 S60416 DNA helicase YGL15
37 87.5 4.7 420 2 D69252 L-carnitine dehydr
38 87 4.7 1085 2 AH3171 DNA polymerase III
39 86.5 4.7 483 1 A53595 allantoinase [EC 3
40 86.5 4.7 605 2 E97717 virE4 protein prec
41 85.5 4.6 252 2 AC2861 conserved hypothet
42 85.5 4.6 276 2 B97638 hypothetical prote
43 85 4.6 396 2 A84556 hypothetical prote
44 85 4.6 718 2 G70978 probable copper-tr
45 85 4.6 974 2 A72012 metalloproteinase,

ALIGNMENTS

RESULT 1
A56098
alpha(1,2)fucosyltransferase Sec2, long form - human
C:Species: Homo sapiens (man)
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 21-Jul-2000
C:Accession: A56098
R:Kelly, R.J.; Rouquier, S.; Giorgi, D.; Lennon, G.G.; Lowe, J.B.
J. Biol. Chem. 270, 4640-4649, 1995
A:Title: Sequence and expression of a candidate for the human Secretor blood group al
tes with the non-Secretor phenotype.
A:Reference number: A56098; MUID:95181460; PMID:7876235
A:Accession: A56098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <KEL>
A:Cross-references: GB:U17894; NID:G687618; PIDN:AAC24453.1; PID:G687619
C:Genetics:
A:Gene: GDB:FUT2; SE
A:Cross-references: GDB:120619; OMIM:182100
A:Map position: 19q13.3-19q13.3
C:Keywords: glycoprotein; Golgi apparatus; transmembrane protein;

Query Match	99.1%;	Score 1832.5;	DB 2;	Length 343;
Best Local Similarity	99.7%;	Pred. No. 5.7e-149;		
Matches 343;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
Qy	1	MLVVQMPFSPMAHILFVFTVSTIFHVQQR	LAKIQAMWELPVQIPVLASTSKALGPSOL	60
Db	1	MLVVQMPFSPMAHILFVFTVSTIFHVQQR	LAKIQAMWELPVQIPVLASTSKALGPSOL	60
Qy	61	RGWWTINATGRIGNQMGVEYATLYALAKMNGRPAFTPAQMHSTLAFIFRITLPLVLSATAS	120	
Db	61	RGWWTINATGRIGNQMGVEYATLYALAKMNGRPAFTPAQMHSTLAFIFRITLPLVLSATAS	120	
Qy	121	RTPWQNYHLNDWMEERYRHI	PPGEYVRFTGYPCSWTFYHHLRQELIQEFTLHDHVREAAQ	180
Db	121	RTPWQNYHLNDWMEERYRHI	PPGEYVRFTGYPCSWTFYHHLRQELIQEFTLHDHVREAAQ	179
Qy	181	KFLRGLOVNGSRPGTFVGVHVRGGDYVHYMPKVKGVADRRYLOQALDWFRRARYSSLIF	240	
Db	180	KFLRGLOVNGSRPGTFVGVHVRGGDYVHYMPKVKGVADRRYLOQALDWFRRARYSSLIF	239	
Qy	241	VVTSNGMACNRNIDTSHGDVVFAGDIGSGPAKDFALLTCQNHITMTTGTGGINAAAYLT	300	
Db	240	VVTSNGMACNRNIDTSHGDVVFAGDIGSGPAKDFALLTCQNHITMTTGTGGINAAAYLT	299	
Qy	301	GGDTIYLANITLPDPSPLKIEKPEAAFLPEWTGIAADLSPLLKH	344	
Db	300	GGDTIYLANITLPDPSPLKIFKPEAAFLPEWTGIAADLSPLLKH	343	

RESULT 2

RESULT 2

B56392
beta-galactoside alphasubscripted alpha,2-fucosyltransferase II - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: B56392

R;Hitoshi S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphasubscripted alpha,2-fucosyltransferase II - rabbit
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: B56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-354 <HIT>
A:Cross-references: GB:X80225; NID:g854356; PIDN:CAA56512.1; PID:g854357
C:Keywords: transmembrane protein

Query Match 74.4%; Score 1375.5; DB 2; Length 354;
Best Local Similarity 76.4%; Pred. No. 7e-110;
Matches 253; Conservative 26; Mismatches 41; Indels 11; Gaps 3;

QY 14 HFLIFVFTSTIHFVOQLAKIQAMWELPQIPVLASTSKALGPSQL--RGMWTINAIGR 71
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 29 YLFTIFVSTVFCHORLALVPAPWAYSARVVV-----PGHLPRGMMTINAMGR 80

QY 72 LGNOMGEYATLYALAKNGRPATFPAQMHSITLAPIFRITLPVLHSATAKSPWNHYLN 131
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 81 LGNOMGEYATLYALAKNGRFAYIPAQMHSITLAPIFRISPLVLSHTASRVPMQNHYLN 140

QY 132 WMEERYRHPPGEVVRFTGYPCSWTFVHHRLQEILQEFTHLDHVREAAQFLSLGVNS 191
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 141 WMEERYRHFPV-PFYRTLTPYCPSWTFVHHRLHEILLREFTHLDHVREAAQAFGLRLVNS 199

QY 192 RPTGFVGHVHRGDYVHMVKWKGVVDARYQQALLDNFRARYSSLIFFVTSGNAWCNR 251
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 200 RPSTFGCVHVRGDYVRVMQVWKGVVDARGYLEQALEDNFRAPTAPPVFTSGNAWCNR 259

QY 252 ENDTSHGDVVFDAGDIEGSPADEFLLTCNHNTIMTGTFGWAAYLTTGGDIYYLANYT 311
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 260 ENIDASRGDVVFAGNLGEGSPADEFLLTCNHVTMTGTGTFGWAAYLTGGDVIYLANYT 319

QY 312 LPDSPFLKFKPEAAFLPEWNTGTAADLSPLL 342
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 320 APDSPFHLVFKPEAEFLPEWGIIANMGRL 350

RESULT 3
A56392
beta-galactoside alphasubscripted alpha,2-fucosyltransferase I - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: A56392

R;Hitoshi, S.; Kusunoki, S.; Kanazawa, I.; Tsuji, S.
J. Biol. Chem. 270, 8844-8850, 1995
A:Title: Molecular cloning and expression of two types of rabbit beta-galactoside alphasubscripted alpha,2-fucosyltransferase I - rabbit
A:Reference number: A56392; MUID:95238380; PMID:7721792
A:Accession: A56392
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <HIT>
A:Cross-references: GB:X80226; NID:g854354; PIDN:CAA56513.1; PID:g854355
C:Keywords: transmembrane protein

Query Match 59.1%; Score 1093.5; DB 2; Length 373;
Best Local Similarity 67.0%; Pred. No. 9.6e-86;
Matches 203; Conservative 38; Mismatches 59; Indels 3; Gaps 2;

QY 41 LPVQIPVLASTSKALGPSQLRGMWTTINAIRGLNQMGYEATLYALAKMNGRPAPIAQMH 100
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 58 LPTYSPASNASSACGRAEAPSGIWTHPDGRFGNQMGQYATLLAQLNGERRAFILPAMH 117

QY 101 STLAIPRTPLPVLSHATSATSRIPQNHYLNDWMEEYVRRHPGPVEFTGYPCSWTIFYHH 160
 : :: |||||:: |||||:: | : | : |||||:: ||
Db 118 AALAPVPRTLPLVLAPEVNRRITSWKOLLHDWMSEYSRL-EDPFLKFTGFCPSWTIFFHH 176

A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially e
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46494
A:Molecule type: mRNA
A:Residues: 1-159 <PIA>
A:Cross-references: EMBL:L26010; NID:9414816; PIDN:AAB41514.1; PID:9414817
C:Genetics:
A:Gene: FTB
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 33.7%; Score 623; DB 2; Length 159;
Best Local Similarity 88.5%; Pred. No. 5.3e-46;
Matches 116; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 213 VKKGVVADRRYLQALDWFARYSSLIFFVTSNGMAWCRENIDTSHGDVVFAGDGIEGSP 272
Db 2 VKKGVVADRGYLEKALDWFARYSSPVFVTSNGMAWCRENINASRGDGVFAGDGIEGSP 61
Qy 273 AKDFALLTQCNTHTMTTGTGFIWAAYLTGSPITLYANLYTLPDPSPLKIFKPEAAFLPEWT 332
Db 62 AKDFALLTQCNTHTMTTGTGFIWAAYLAGGDTIYLYANLYTLPDPSPLKIFKPEAAFLPEW 121
Qy 333 GAAADLSPLK 343
Db 122 GIPADLSPLK 132

RESULT 6
S51582
galactoside 2-alpha-L-fucosyltransferase (EC 2.4.1.69) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1995 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C:Accession: S51582; S46493
R:Piau J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Submitted to the EMBL Data Library, November 1993
A:Reference number: S51582
A:Accession: S51582
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-142 <PIA>
A:Cross-references: EMBL:L26009; NID:9414814; PIDN:AAB41514.1; PID:9554438
R:Piau, J.P.; Labarriere, N.; Dabouis, G.; Denis, M.G.
Biochem. J. 300, 623-626, 1994
A:Title: Evidence for two distinct alpha-(1,2)-fucosyltransferase genes differentially e
A:Reference number: S46493; MUID:94280382; PMID:8010942
A:Accession: S46493
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 9-135 <PI2>
A:Cross-references: EMBL:L26009
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 32.1%; Score 593; DB 2; Length 142;
Best Local Similarity 74.6%; Pred. No. 1.7e-43;
Matches 106; Conservative 16; Mismatches 18; Indels 2; Gaps 1;

Qy 152 PCSWTFYHHLROFILOEFTLHDHVRERAAQKFLRGQV--GSRPGTFVGVHVRGDIYVHV 209
Db 1 PCSWTFYHHLRQIRREFTLHDHVRERAAQKFLRGQV--GSRPGTFVGVHVRGDIYVHV 60
Qy 210 MPKVKGVVADRRYLQALDWFARYSSLIFFVTSNGMAWCRENIDTSHGDVVFAGDGIE 269
Db 61 MPKVKGVVADRRYLQALDWFARYSSLIFFVTSNGMAWCRENIDTSHGDVVFAGDGIE 120
Qy 270 GSPAKDFALLTQCNTHTMTTGT 291
Db 121 GTPGKDFALLTQCNTHTMTTGT 142

RESULT 7
T44328
hypothetical protein wblA [imported] - Vibrio cholerae
C:Species: Vibrio cholerae

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 28-Jul-2000
C:Accession: T44328
R:Yamasaki, S.; Shimizu, T.; Hoshino, K.; Ho, S.T.; Shimada, T.; Nair, G.B.; Takeda,
Gene 237, 321-332, 1999
A:Title: The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are cl
A:Reference number: 222749; MUID:99453293; PMID:10521656
A:Accession: T44328
A:Status: preliminary; translated from GE/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-281 <YAM>
A:Cross-references: EMBL:AB012957; NID:94115688; PIDN:BAA33632.1; PID:93721682
A:Experimental source: strain O22
C:Genetics:
A:Note: wblA
C:Superfamily: *Vibrio cholerae* hypothetical protein wblA

Query Match 10.5%; Score 195; DB 2; Length 281;
Best Local Similarity 26.6%; Pred. No. 4.4e-09;
Matches 77; Conservative 46; Mismatches 104; Indels 62; Gaps 14;

Qy 70 GRIGNQMGAYATLYALAKMNGRP-----AFTPAQMHSYL-APIFRITLPVLH----- 115
Db 8 GGLGNLFQYAVGRAIAIQYGVPLKLDYSAYKNYKLNHGYRLDQFNADIANEDIEIHL 67
Qy 116 -----SATASRIPWQNYHLNDWMEEEVRI-----PPGEYVFTGYPGCSWTFYHHL 161
Db 68 KGSSNRLSRILRLRGW--LKKNTYYAEKQRTYDVSVFMQAPRYL--DGYWQNEQYFSOI 123
Qy 162 ROEILQEFTHDVRERAAQKFLRGQVNGSRPGTFVGVHVRGDIYVHVMPKVKGVVADR 221
Db 124 RAVLQELWPNQPLINAAQHQIKQOQTHA-----VSHVRGDIYLN-HPEIG---VLDI 174
Qy 222 RYLOQALDWFARYSSLIFFVTSNGMAWCREN---IDTSHGDVVFAGDGIEGSPAKDFAL 278
Db 175 DYYKRAVDYIKKIEAPVFEVFNDAWCKDNFNEDSP---VFIED--TQTEIDDLML 228
Qy 279 LTQCNTHTMTTGTGFIWAAYLTGG-DTYLYANT-----LPDS 315
Db 229 MCQCQHNIVANSFSSWAALNSRDYKIVAPKTWMAENPKYKWKVPDS 277

RESULT 8
T20745
hypothetical protein F11A5.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T20745
R:Gardner, A.
Submitted to the EMBL Data Library, March 1997
A:Reference number: Z19319
A:Accession: T20745
A:Status: preliminary; translated from GE/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-363 <WIL>
A:Cross-references: EMBL:292830; PIDN:CAB07352.1; GSPDB:GN00023; CESP:F11A5.5
A:Experimental source: clone F11A5
C:Genetics:
A:Gene: CESP:F11A5.5
A:Map position: 5
A:Introns: 42/1; 55/2; 136/3; 175/3; 210/2; 243/1; 299/1; 330/3
C:Superfamily: *Caenorhabditis elegans* hypothetical protein C14C6.3

Query Match 9.8%; Score 181.5; DB 2; Length 363;
Best Local Similarity 23.3%; Pred. No. 8.6e-08;
Matches 87; Conservative 54; Mismatches 116; Indels 117; Gaps 20;

Qy 16 ILFVFTVSTIFHVQORLAKIQAMWEL-PVQIPVLASTKALGPSQLRGWMTNAI----- 59
Db 33 LEWMILVYSIF-----LIRAMHEEDPEVP-----YPPAQINFTSRKYISSNYA 77
Qy 70 --GRIGNQMGAYATLYALAKMNGR-PAFIPAQM-----STLAPIFRITLPVLHSAT 118
Db 78 SSSRLGNHLFELASVLSISRELQRFVTFIENCYHEKMWEDSNLTLPGLMNHFLINGSV 137

QY 119 ASRIPQNVH-----LNDWMEERYRHIPPGEYVVRFTGYPCSWTFYHHLRQBILOE 166
 Db 138 PSSVKRYKHQKCCFFDSSLNDNYDEYHLTGTHYO-----SWKYFESHMRNELIGY 190
 QY 169 FTLHDHVRREEAQKFLRGLOVNGSRGTFYGV-VHVRGDDYVHVMPKVGK-VADRRYLAQ 226
 Db 191 LKTTENTYMDLPK-----SGENTFITCVHVRGDFLRV-----CFHVADENFIRS 235
 QY 227 ALDWFPRYSSLIFFVTSNGMAWCHRENDTSHGDDVVFAGDGE----- 269
 Db 236 SUNLISRQVAK-----RANTAT-----VFFGDDYEFMDSLRNR'SKINAEVS 277
 QY 270 -GSPAKDFALL-TCQNHITMTI--GTGFWAAAYLGGDTIYLANVTLDPSPFLK----- 319
 Db 278 QNSPADLLYAKSNCDDVLLTAHSTFGWGMGYFSKGRVY---YT--DIQFTKDWILET 332
 QY 320 -IFKPEAAFLPEWT 332
 Db 333 GEFISEDYILPHWT 346

RESULT 9
 H87911
 protein B0205.4 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: H87911
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A/Reference number: A75000; MURD:99069613; PMID:9851916
 A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
 A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A/Accession: H87911
 A>Status: preliminary
 A:Molecule type: DNA
 A/Residues: 1-443 <STO>
 A/Cross-references: GB:chr.I; PIDN:AAC16988.1; PID:g3150470; GSPDB:GN00019; CESP:B0205.4
 A/Note: contains weak similarity to fucosyltransferases
 C:Genetics:
 A:Gene: B0205.4
 A:Map position: 1

Query Match 9.5%; Score 176; DB 2; Length 443;
 Best Local Similarity 20.4%; Pred. No. 3.3e-07;
 Matches 89; Conservative 63; Mismatches 145; Indels 140; Gaps 17;

QY 15 FILEVFTVS-----TIFHVQORLAKI-QAMWELPV-----QIPVLA 49
 Db 29 YILFLENLSKTRLTPEERPATYMKKNVYQSGNEAILDNCPW--PISNVENVINRVSLE 86
 QY 50 STSKA-----LGPSQLRGWMTINAGRLGNMGCEYATLYALAKMNGRPAPFIPAOMHSTLA 104
 Db 87 NSCKRIFSDFGYSQ-----GIGNLMPQVAGLLSIARETSILLIPST--TTLR 133
 QY 105 PIF-----RTTLVPLHSATASRIIPWON---VHLNDWMEERYRHIPPGEY 145
 Db 134 RAFDEFETTNDSTQFVGEDLSROLAEDLNASKITLTSCCAYRNLSILFNDRSRIIE--- 189
 QY 145 VRFEGYPCSWTYHHLRQBILOE-FTLHDHVR----- 177
 Db 190 -RIDGYFQNFYFHPDSOKIVKLTFFMDPVKRKVSFWNIYVNIHTPHNRKPKSTVS 248
 QY 178 -----EAKQFLR---GLQVNGSR-----PGTFVGVHVR 202
 Db 249 IFSVTFVQLRVDFLENVGISTVTRNARIETNVANDQALPEEDAFAKTMVGVHLR 308
 QY 203 RGDYVHV--MPKWKGVADRRYLOALDWFPRYSSLIFFVTSNGMAWCHRENDTSHGD 260
 Db 309 HGMDSMNSRNIHGHVDTPIEYKRAIQOISKIYENAFITCSNDVAVARNLKGKET 368
 QY 261 VYFAGDIEGSPAKDFALLTQCNHITMTITGTFGIWAAYL---TGDDTIYLANVTLDPSPFF 317

Db 369 LHEFF---CPGPREVDMAILKSCDSVLIISTGTFGWNSAYLNNASPDVYIYKHWPAPGSVM 425
 QY 318 LKIFKPEAAFLPEWTGI 334
 Db 426 EXMINKTEYFLKSWTAL 442

RESULT 10
 T32309
 hypothetical protein F31F4.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T32309
 R:Blanchard, M.; Kramer, J.; Elliott, G.; Twyman, B.
 submitted to the EMBL Data Library, September 1997
 A/Description: The sequence of C. elegans cosmid F31F4.
 A/Reference number: Z21149
 A/Accession: T32309
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A/Residues: 1-395 <BLA>
 A/Cross-references: EMBL:AF024503; PIDN:AAB70387.1; GSPDB:GN00023; CESP:F31F4.11
 A/Experimental source: strain Bristol N2; clone F31F4
 C:Genetics:
 A:Gene: CESP:F31F4.11
 A:Map position: 5
 A/Introns: 22/3; 98/2; 179/3; 214/2; 248/1; 283/3; 329/1; 360/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 8.4%; Score 155; DB 2; Length 395;
 Best Local Similarity 20.4%; Pred. No. 1.8e-05;
 Matches 77; Conservative 47; Mismatches 110; Indels 144; Gaps 18;

QY 18 FVFTVSTIFHVQORLAKIQAMWELPVQIP---VLASTSKALGSPQLRGMTWITNIGRLGN 74
 Db 80 YFSAETLF-----SKNSNFWLSGIPYPPALINFTSSQVLSNLSAPC-----QLGN 127
 QY 75 QMGVATLYALAK-MNGRPA-FIPAQMHSILAPITRTPVLVHSATASRIPWQNYHLNDW 132
 Db 128 HIFELASLYGLSKHLNRTPAFFESGYHKNMLDSIRSTMPGL----- 169
 QY 133 MEEYRHIPPGEYVVRFTGYPCSWTFYHHLRQBILOEFTLHDHVRREEAQKFLRGLOVNGSR 192
 Db 170 -----IGKYAIFDG--SSWKYFPNREELI-----SFLNTSVQNF--GI-LPISN 209
 QY 193 PGTFVG-VHVRGDDYVHVMPKVGKGVVADRRYLOALDWFPRYSSLIFFVTSNGMAWCR 251
 Db 210 NNTHVSCVHSRRGDFEVM-----NFYATDPKMKNAVKELE-----K 247
 QY 252 ENIDTSHGDDVVFAGDGI-----EGSPAKDFALLT-QCNHT 285
 Db 248 ENYSKKNQKIVLFGDDFKPMRNLFSEAKVSTDAHESVEYYISONSADIDFLYSAYNCDEV 307
 QY 286 IMP-----IGTFGWAAYLTGDDTIYLA-----NYTLDP 314
 Db 308 LITGNAHQPTSMHVFANFNLLDTPRSTFGWNLGYFGSKGVYIYLDIKHARDVYDNGEIAE 367
 QY 315 SPPLKIFKPEAAFLPEWT 332
 Db 368 NDF-----FLPHWT 376

RESULT 11
 T31916
 hypothetical protein C17A2.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T31916
 R:Sammons, L.; Wohldmann, P.
 submitted to the EMBL Data Library, July 1997
 A/Description: The sequence of C. elegans cosmid C17A2.
 A/Reference number: Z21098

125334
hypothetical protein T26H5.8 - *Caenorhabditis elegans*

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T25334; T26600
 R:Gardner, A.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z20017
 A:Accession: T25334
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-392 <WIL>
 A:Cross-references: EMBL:282056; PIDN:CAB04857.1; GSPDB:GN00023; CESP:T26H5.8
 A:Experimental source: clone T26H5
 R:Wall, M.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20243
 A:Accession: T26600
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-392 <WIL>
 A:Cross-references: EMBL:AL032620; PIDN:CAA21488.1; CESP:T26H5.8
 A:Experimental source: clone Y3E3A
 C:Genetics:
 A:Gene: CESP:T26H5.8
 A:Map position: 5
 A:Introns: 7/1; 32/3; 55/1; 109/3; 148/3; 187/2; 222/1; 263/1; 333/1; 363/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C14C6.3

Query Match 7.4%; Score 137.5; DB 2; Length 392;
 Best Local Similarity 21.0%; Pred. No. 0.00055;
 Matches 72; Conservative 56; Mismatches 114; Indels 101; Gaps 18;

Qy	72	LGQMGEYATLYALAKMGRPAFIPA-----QMHSLAPITF-----RITLPVLHSA	119
Db	54	LGRLFELISLIGIASTLQRRVAVINATPNVNSHLLQTMQLPFLKVEQFELRVIPESV	113
Qy	120	SR--IPWQNYHLND---WMEEYRH-IPPEYVFTGYPGCSWTFFYHHLROEILQEF	172
Db	114	ARQINWSRCIFDDPAKVLDISNOHVLDGHYQ-----SFKIFHIRKIREWMAKP	167
Qy	173	DHVEEAQKFLRGLOVNGSRGRTGVGVHVRGDY-----VHYMPKVKGVADRRYLOAL	228
Db	168	ELAKSAEKLI-----PFSVKDNFICPHIRRGDTDGLH-QP-----SDPTFTRA	215
Qy	229	DWFRARYSS-----LIFVVTSGMA-----WCRENIDTSH--GDVVFAGDG-	267
Db	216	DFLVEKQKQIRRHVTVVVFGNDMFAHTVFEKGSFFFECKKTYEILHRVGNISFALNP	275
Qy	268	-----IEGSPA-----KDFALL-----TCNHTMT--I	289
Db	276	ATALNYLTPEYSPYDVRSRNFALATLKHQVILTPSTPEIDLAFSKTFCDALFTGPS	335
Qy	290	GTFGIWAAYLTGGDT-IYLANYTLPDSFFLKIFKPEAAFLPEW	331
Db	336	STFGWNL SylAKSSAKVYIRDIETLKGVDNMKVEDFYPPW	378

RESULT 15
 T02798
 hypothetical protein L549.10 [imported] - Leishmania major (strain Friedlin)
 C:Species: Leishmania major
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C:Accession: B81456; T02798
 R:Myler, P.J.; Audleman, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A:Reference number: B81455; MUID:99178987; PMID:10077609
 A:Accession: B81456
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <PVL>
 A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24622.1; PID:g2978459; GSPDB:GN00
 A:Experimental source: strain MHOM/IL/81/Friedlin
 C:Genetics:

A:Gene: L549.10
 A:Map position: 1

Query Match 7.1%; Score 131.5; DB 2; Length 348;
 Best Local Similarity 18.1%; Pred. No. 0.0015;
 Matches 67; Conservative 48; Mismatches 125; Indels 131; Gaps 11;

Qy	54	ALGPSQLRGMTTINATGRGNQGEYATLYALAKMGRPAFIPA-----	97
Db	13	AATPSPSKARLTNTIYGGMGNLFLVANLLATAHRTGIPAYLEAVPSSSAEDPRFTYND	72
Qy	98	-----QMHSLAPITFRTLPVL-----	116
Db	73	TLFRDLGHYGVQMCASSPPLPVVTPETPRVKVQLDAQRPCVFMHIGFTQSDAFDDHP	132
Qy	117	ATASRTP---WQN-----YHLNDWMEFEYRHIPPEYVFTGYPGCSWTFFYHHLROEIL	166
Db	133	INSSVTPRELWDSAGKHLATYYGGSGCHVALHVRGDTYRFT-----DIF	178
Qy	167	QEFTHLDHYREEAQKFLRGLOVN-----GSRPGTFVGVHVRGEY	206
Db	179	EQLDVVEYYDAARVQLLGLLQSPSLLRQQSPVDSQAQPTAGAPP-----	225
Qy	207	VHMPKVKGVADRRYLOALDWPRARYSSLIYV-----VTSNGMAWCRENIDTSHGDVV	262
Db	225	--LLPLHLVFCDEBERFGRTVGYFRTKYGAVMVSVCALTEASAIARPATSSPSVFI	283
Qy	263	FAGDIEGSPAKDFALLTQCNIHITIGTGIWAAYLTGGDTIYLANYTLPDSPFLKIFK	322
Db	284	MFRDYLE-----LLMMSOCNDVVMANSTFSWGAYL-----NRVSLRVTAPSRNFVKD	334
Qy	323	PEA--AFLPEW	331
Db	335	FASNHLXCPGW	345

Search completed: May 27, 2003, 15:08:20
 Job time: 14.4655 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:53:11 ; Search time 10.0683 Seconds
(without alignments)

1417.108 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVYMPFSPFMAHFLVFE.....AAFLPEWTGIAADLSPLIKH 344

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	1832.5	99.1	343	FUT2_HUMAN	Q10981 h galactosi
2	1528	82.6	340	FUT2_PIG	Q10982 sus scrofa
3	1510	81.6	344	FUT2_BOVIN	Q28113 bos taurus
4	1375.5	74.4	354	FUT2_RABIT	Q10983 oryctolagus
5	1299.5	70.2	368	FUT2_MOUSE	P97353 mus musculus
6	1100.5	59.5	376	FUT1_RAT	Q10980 rattus norv
7	1093.5	59.1	373	FUT1_RABIT	Q10979 oryctolagus
8	1083.5	58.6	376	FUT1_MOUSE	Q09160 mus musculus
9	1080.5	58.4	365	FUT1_HUMAN	P19326 homo sapien
10	1079.5	58.4	365	FUT1_PIG	Q29043 sus scrofa
11	623	33.7	159	FUT2_RAT	Q10984 rattus norv
12	155	8.4	365	YK07_CAEEL	P34302 caenorhabdi
13	52	5.0	575	FUT8_PIG	P79282 s alpha-(1,
14	90	4.9	575	FUT8_HUMAN	Q9bys5 h alpha-(1,
15	90	4.9	575	FUT8_MOUSE	Q9wt52 m alpha-(1,
16	89	4.8	1489	YGP0_YEAST	P53115 saccharomyc
17	86.5	4.7	484	ALN_RANCA	P40757 rana catesb
18	85	4.6	323	MIAA_RALSO	Q8xb0 ralstonia s
19	85	4.6	718	CTPC_MYCTU	P96875 mycobacteri
20	84.5	4.6	630	S6A4_CAVPO	O35899 cavia porce
21	84	4.5	433	CBPA_ANOGA	O02350 anopheles g
22	83.5	4.5	660	FHUB_ECOLI	P06972 escherichia
23	83	4.5	328	GLK_NEIMA	Q9jx33 neisseria m
24	82.5	4.5	1048	SBCC_ECOLI	P13458 escherichia
25	82	4.4	575	FUT8_BOVIN	Q9n0w2 b alpha-(1,
26	82	4.4	1211	ATC6_SCHPO	Q14072 schizosacch
27	82	4.4	1928	LPH_RAT	Q02401 rattus norv
28	81.5	4.4	476	NOEA_RHIME	O52892 rhizobium m
29	81.5	4.4	537	FUT6_ARATH	Q9xi80 arabidopsis
30	81.5	4.4	630	S6A4_MOUSE	O60857 mus musculus
31	81.5	4.4	630	S6A4_RAT	P31652 rattus norv
32	81.5	4.4	1279	SCAP_HUMAN	Q12770 homo sapien
33	80.5	4.4	874	BCSA_SALTY	Q93in2 salmoneilla

34	80	4.3	990	1	KGPL_PICPA	Q92448 picidia past
35	79.5	4.3	353	1	FES_YEREN	Q58555 yersinia en
36	79	4.3	454	1	NOE2_HUMAN	Q95897 homo sapien
37	79	4.3	776	1	PSAB_HETTR	Q9xgv2 heterocapsa
38	79	4.3	1461	1	TOP2_CANAL	P87078 candida alb
39	78.5	4.2	645	1	L1PL_PHOLU	P40601 photorhabdu
40	78	4.2	470	1	PABB_BACSU	P28820 bacillus su
41	78	4.2	890	1	POL2_BAYMG	Q01365 bailey yell
42	78	4.2	963	1	TRES_THETH	O08458 thermus the
43	78	4.2	2294	1	YCF2_ARATH	P56786 arabidopsis
44	78	4.2	2479	1	POLN_RRVN	P13887 ross river
45	77.5	4.2	395	1	HMPA_ERWCH	Q47266 erwinia chr

ALIGNMENTS

RESULT 1						
FUT2_HUMAN						
ID	FUT2_HUMAN	STANDARD;	PRT;	343 AA.		
AC	Q10981;					
DT	01-OCT-1996 (Rel. 34, Created)					
DT	01-OCT-1996 (Rel. 34, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-					
DE	D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)					
DE	(Fucosyltransferase 2) (Secretor blood group alpha-2-					
DE	fucosyltransferase) (Secretor factor) (Se) (SE2).					
GN	FUT2 OR SEC2.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID:9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=95181460; PubMed=7876235;					
RA	Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;					
RT	"Sequence and expression of a candidate for the human Secretor blood					
RT	group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an					
RT	enzyme-inactivating nonsense mutation commonly correlates with the					
RT	non-secretor phenotype."					
RL	J. Biol. Chem. 270:4640-4649(1995).					
RN	[2]					
RP	SEQUENCE FROM N.A., AND VARIANT PHE-140.					
RX	MEDLINE=96199252; PubMed=8621666;					
RA	Kudo T., Iwasaki H., Nishihara S., Shinya N., Ando T., Narimatsu I.,					
RA	Narimatsu H.;					
RT	"Molecular genetic analysis of the human Lewis histo-blood group					
RT	system. II. Secretor gene inactivation by a novel single missense					
RT	mutation A385T in Japanese nonsecretor individuals."					
RL	J. Biol. Chem. 271:9830-9837(1996).					
RN	[3]					
RP	SEQUENCE FROM N.A., AND VARIANT PHE-140.					
RX	MEDLINE=97363210; PubMed=9219535;					
RA	Koda Y., Soejima M., Wang B., Kimura H.;					
RT	"Structure and expression of the gene encoding secretor-type					
RT	galectoside 2-alpha-L-fucosyltransferase_(FUT2)."					
RL	Eur. J. Biochem. 246:750-755(1997).					
RN	[4]					
RP	VARIANTS VAL-25; CYS-138 AND ASN-172.					
RX	MEDLINE=98431007; PubMed=9760207;					
RA	Liu Y., Koda Y., Soejima M., Pang H., Schlaphoff T., du Toit E.D.,					
RA	Kimura H.;					
RT	"Extensive polymorphism of the FUT2 gene in an African (Xhosa)					
RT	population of South Africa."					
RL	Hum. Genet. 103:204-210(1998).					
CC	!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA					
CC	{(1,2)GALBETA-} CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL					
CC	SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN					
CC	SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR					
CC	SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.					
CC	!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +					
CC	alpha-L-fucosyl-1,2-beta-D-galactosyl-R.					

[illegible][illegible]

```

RESULT 4
FUT2_LABIT
ID      FUT2_LABIT      STANDARD;      PRIT;      354 AA.
AC      Q10383;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (Secretor blood
DE      group alpha-2-fucosyltransferase) (GDP-1-fucose:beta-D-
DE      galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)F T 2)
DE      (Fucosyltransferase 2).
DE      FUT2 OR SECL CR RFT-II.
GN      Oryctolagus cuniculus (Rabbit).
OS      Oryctolagus cuniculus (Rabbit).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX      NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
RP      TISSUE=Brain;
RC      MEDLINE=95238380; PubMed=7721792;
RX      Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RA      "Molecular cloning and expression of two types of rabbit beta-
RT      galactoside alpha 1,2-fucosyltransferase.";
RL      J. Biol. Chem. 270:8844-8850(1995).
CC      !- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC      ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC      SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A'AND B ANTIGEN
CC      SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC      SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC      !- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +

```

```
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: SALIVARY AND LACTATING MAMMARY GLANDS.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80225; CAA56512.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 44 354 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 40035 MW; 1E2B831F9DA6CCB4 CRC64;
SQ SEQUENCE 354 AA; 40035 MW; 1E2B831F9DA6CCB4 CRC64;

Query Match 74.4%; Score 1375.5; DB 1; Length 354;
Best Local Similarity 76.4%; Pred. No. 1.4e-110;
Matches 253; Conservative 26; Mismatches 41; Indels 11; Gaps 3;

Qy 14 HFILEFTVSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQL--RGMTWTAIGR 71
Db 29 YLFTTFIVSTVFHCHORLALVPAPAYSAKVVV-----PGHLPREGMTWTAIGR 80
Qy 72 LGNQMGVATYALAKMNGRPAFIPAQMHSHTLAPFRITPLVLSHATSRIPWQNYHLND 131
Db 81 LGNQMGVATYALAKENGRPAFIPAQMHSHTLAPFRISLPLVLSHSTASRYWQNYHLND 140
Qy 132 WNEEYRHIPPGEYVRFYPCSWTFYHLLRQELIQLPFTLHDHVREREAQKFLRGQVNGS 191
Db 141 WNEEYRHIPPYVRLTYGPCSWTFYHLLRQELIQLPFTLHDHVREREAQKFLRGQVNGS 199
Qy 132 PPGTFVGVHVRGDDYVHVPKWKGVVADRYVLOQALDWERARYSSLIFFVVTSGMAWCR 251
Db 200 RPSTFVGVHVRGDDYVHVPKWKGVVADRYVLOQALDWERARYSSLIFFVVTSGMAWCR 259
Qy 252 ENIDTSHGDVVFAGDIEGSPAKDFALLTQCNRHMTITGTCGIAAAYLTGGDTIYLANYT 311
Db 260 ENIDASRGDVFVAGNGLBGSFPAKDFALLTQCNRHMTITGTCGIAAAYLTGGDTIYLANYT 319
Qy 312 LPDSPFLKIFKPEAFLEPWGTGIAADUSPLL 342
Db 320 APDSPFLHVFKEAFLEPWGTGIANMGRAL 350

RESULT 5
FUT2_MOUSE
ID FUT2_MOUSE STANDARD; PRT; 368 AA.
AC P97353;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 2 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 2) (Alpha(1,2)Ft 2)
DE (Fucosyltransferase 2) (FUT-III).
GN FUT2 OR SECL.
OS Mus musculus (Mouse).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Htoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of a mouse GDP-L-Fucose: beta-D-
RT galactoside 2-alpha-L-Fucosyltransferase."
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Gastrointestinal tract;
RA Lin B., Hayashi Y., Saito M., Sakakihara Y., Yanagisawa M.,
RA Iwamori M.;
RT "Molecular cloning and expression of a GDP-L-fucose:beta-D-galactoside
RT 2-alpha-L-fucosyltransferase in murine gastrointestinal tract."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND SECL) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y09882; CAA71008.1; -.
DR EMBL; AF113532; AAD25351.1; -.
DR MGD; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 42 368 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 368 AA; 41464 MW; 4093E853EB37303B CRC64;

Query Match 70.2%; Score 1299.5; DB 1; Length 368;
Best Local Similarity 73.6%; Pred. No. 4.8e-104;
Matches 245; Conservative 30; Mismatches 43; Indels 15; Gaps 4;

Qy 11 PWAHFLF--VFTVSTFHFVQORLAKIQAMWELPQVIPVLASTSKALGPSQ---LRGMWT 65
Db 22 PLSTFVLFVIFVSTVFHCHRRGLGVAPW-----ASPSLVVFPFRMPREGNET 72
Qy 66 INATGRGLGNQMGVATYALAKMNGRPAFIPAQMHSHTLAPFRITPLVLSHATSRIPWQ 125
Db 73 IRVKGRLGNQMGVATYALAKMNGRPAFIPAQMHSHTLAPFRISLPLVLSHSDAKRIPWQ 132
Qy 126 NYHLNDWMEERYRHIPPGEYVRFYPCSWTFYHLLRQELIQLPFTLHDHVREREAQKFLRG 185
Db 133 NYHLNDWMEERYRHIPPGEYVRFYPCSWTFYHLLRQELIQLPFTLHDHVREREAQKFLRG 191
Qy 196 LOVNGSRGCTFVGHVVRGDDYVHVPKWKGVVADRYVLOQALDWERARYSSLIFFVVTSGN 245
Db 192 LOVNGSRGCTFVGHVVRGDDYVHVPKWKGVVADRYVLOQALDWERARYSSLIFFVVTSGN 251
Qy 246 GMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNRHMTITGTCGIAAAYLTGGDTI 305
Db 246 GMAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNRHMTITGTCGIAAAYLTGGDTI 305
```

```
Db 252 DMWCRKSITASRGDNVAFAGNLOGSPAKDIALMQCNHTVITLCTFGIWAAYLTGGDTV 311
QY 306 YLANETLPDSFELKIFKPEAAFLPEWTGIAADL 338
Db 312 YLANETQNSPFHTVFKPEAAFLPEWVGIAADL 344

RESULT 6
FUT1_RAT
ID FUT1_RAT STANDARD; PRT; 376 AA.
AC Q10980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1 OR FTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon cancer;
RA Soejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-alpha-L-
RT fucosyltransferase genes.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 169-310 FROM N.A.
RC STRAIN=BDIX;
RX MEDLINE=94280382; PubMed=8010942;
RA Piau J.-P., Labarriere N., Dabouis G., Denis M.G.;
RT "Evidence for two distinct alpha(1,2)-fucosyltransferase genes
RT differentially expressed throughout the rat colon.";
RL Biochem. J. 300:623-626(1994).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB015637; BAA31130.1; -
CC EMBL; AB006137; BAA21741.1; -
CC EMBL; L26009; AAB41514.1; -
CC Inter-Pro: IPR002516; GT-11.
CC Pfam: PF01531; Glyco.transf.11; 1.
CC Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 12
CC TRANSMEM 13 29
CC CYTOPLASMIC (POTENTIAL).
CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC LUMENAL, CATALYTIC (POTENTIAL).
CC CARBOHYD 30 376
CC CARBOHYD 64 64
CC CARBOHYD 302 302
CC CARBOHYD 328 328
CC CARBOHYD 328 328
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC C350C737C758B7F8 CRC64;
CC SEQUENCE 376 AA; 42416 MW; 42416 MW;
Query Match 59.5%; Score 1100.5; DB 1; Length 376;
```

```
Best Local Similarity 67.0%; Pred. No. 5.7e-87;
Matches 207; Conservative 36; Mismatches 59; Indels 7; Gaps 4;

QY 37 AMWELPQIPVLASTSKALGPSQ--LRGMWTINAGRIGNQMGLEYATLYALAKMKGRPAF 94
Db 55 AMVCLP--YPSNASSGSPSCPEQLSLSGTWTITPGRTGNQMGQYATLLALAQNGRAAF 112
QY 95 IPAQMHSLAPIFRITLPLVLSHATASRIPWQNYHLNDMMEEYRHIIPPGEYVRFYGYPCS 154
Db 113 IOPEMHTTLPVFRISLPVLDPEVDSLTPWQLVLDMMSEYSHL-BDPFKLSGFPSC 171
QY 155 WIFYHHLROEILQEFTLHDHVRERAKFELRGQVN--GSRPGTFVGVHVRGDDYVHVMPK 212
Db 172 WYFFHHLRQIRREFTLHDHLREDQAQLSLGRLGIPAGIRPRTYGVHVRGDDYLEVMFN 231
QY 213 VNKGVVADRYLQOALDWPFRARYSSLIFFVVTSGNWCARENIDTSGDVVFAGDGLGESP 272
Db 232 RNKGVVGDRAYLQKAMDWFRARHKDPIFVVTSGNWCARENIDTSGDVVFAGDGLGESP 291
QY 273 AKDFALLTQCNHTIMTIGTFGFWAAYLTGSDTIYLANYTLPDSPFLKIFKPEAAFLPEWT 332
Db 292 GKDFALLTQCNHTIMTIGTFGFWAAYLAGDVIYLANFLTDPSEFLKIFRPAFLPEWV 351
QY 333 GIAADLSPL 341
Db 352 GINADLSPL 360

RESULT 7
FUT1_RABBIT
ID FUT1_RABBIT STANDARD; PRT; 373 AA.
AC Q10979;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1 OR RFT-1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95238380; PubMed=7721792;
RA Hitoshi S., Kusunoki S., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of two types of rabbit beta-
RT galactoside alpha 1,2-fucosyltransferase.";
RL J. Biol. Chem. 270:8844-8850(1995).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X80226; CAA56513.1; -
CC Inter-Pro: IPR002516; GT_11.
```

```

DR Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 30 373 (POTENTIAL).
FT DOMAIN 30 373 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42098 MW; 0A47A1786231525C CRC64;

Query Match 59.1%; Score 1093.5; DB 1; Length 373;
Best Local Similarity 67.0%; Pred. No. 2.2e-85;
Matches 203; Conservative 38; Mismatches 59; Indels 3; Gaps 2;

Qy 41 LPVQIPVLASTSKALGPSQLRGMTINAIIRGLNQMGQYATLYALAKMNGRPAPIPQMH 100
Db 58 LPVTSASNASSACGRPAAPSGIWTIHPDGRFGNQMGQYATLLAOLNGRAFLIPAMH 117
Qy 101 STLAPEIRITLVLHSATASRIPQNYHLNDWMEERYHIPPGEVVRFTGPGCSWTFYHH 160
Db 118 AALAPVFRITLVLPAPEVNRVTSKQALLHDMSEYSEL-EDPFLKLTGPGCSWTFEHR 176
Qy 161 LROETLOFTLHDHVRERAKFLRGLQV--NGSRPGTFVGVYVRRGDYVHVMPKVKGW 218
Db 177 VREQIRREFTLHDHVRERAKFLRGLRGTGARPTFVGVYVRRGDYVHVMPKVKGW 236
Qy 219 ADRRYLOALQDFRARSLSLIFVTSNGMACRENIDTSHGDVVFAGDGIGSPAKDFAL 278
Db 237 GDRAYLOQAMDWRFAHKEAPIFVTSNGMKWCENIDASRGDVFVAGNGLSSPFAKDFAL 296
Qy 279 LTCNHTMTGTGFWAAYLTGGDTIYLANFTLPDSPLKIFKPEAAFLPWTGIAADL 338
Db 297 LTCNHTMTGTGFWAAYLAGGDTIYLANFTLPDSPLKIFKPEAAFLPWTGIAADL 356
Qy 339 SPL 341
Db 357 SPV 359

RESULT 8
FUT1_MOUSE STANDARD; PRT; 376 AA.
ID FUT1_MOUSE
AC C09160;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
GN FUT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss;
RX MEDLINE=97454449; PubMed=9355741;
RA Domino S.E., Hiraiwa N., Lowe J.B.;
RT "Molecular cloning, chromosomal assignment and tissue-specific expression of a murine alpha(1,2)fucosyltransferase expressed in thymic and epidermal epithelial cells.";
RL Biochem. J. 327:105-115(1997).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP + alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: IN THE ADULT, HIGHLY EXPRESSED IN PANCREAS, TESTIS AND EPIDIDYMIS AND TO A LESSER EXTENT IN THYMUS, LUNG, STOMACH, SMALL INTESTINE, COLON, SPLEEN AND UTERUS. NOT EXPRESSED

```

```

CC IN BRAIN, HEART, SKELETAL MUSCLE, KIDNEY, LIVER AND BONE MARROW.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U90553; AAC53492.1; .
CC MGD: MGI:109375; Fut1.
CC InterPro: IPR002516; GT_11.
CC Pfam: PF01531; Glyco_transf_11; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 9 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 27 376 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 376 AA; 42255 MW; 21FID24CFE204106 CRC64;

Query Match 58.6%; Score 1083.5; DB 1; Length 376;
Best Local Similarity 67.0%; Pred. No. 1.6e-85;
Matches 201; Conservative 33; Mismatches 61; Indels 5; Gaps 3;

Qy 46 PVLASTSKALGPSQ---LRGMWTTINAIIRGLNQMGQYATLYALAKMNGRPAPIPQMHST 102
Db 62 PLOTNSGPSQSPCEQSSSLSGTWTITPGGRFGNQMGQYATLLAOLNGRAQFIQPEMHA 121
Qy 103 LAPTRITLVLHSATASRIPQNYHLNDWMEERYHIPPGEVVRFTGPGCSWTFYHHR 162
Db 122 LAPVERISLPVLDPEDVSLTPQWHLVLDHWMSEYSHL-EDPFLKLSGFPSCWTFEHLR 180
Qy 163 OEILQEFTHDHRERAKFLRGLQVNSRPG-TFVGVRVRRGDYVHVMPKVKGWADR 221
Db 181 EQRREFTLHNLHREGAOYLLSLRIGPASPARTFVGVYVRRGDYVHVMPKVKGWVDR 240
Qy 222 RYLOALQDFRARSLSLIFVTSNGMACRENIDTSHGDVVFAGDGIGSPAKDFALLTQ 281
Db 241 AYLOQAMDWRFAHKEAPIFVTSNGMKWCENIDTSHGDVVFAGDGIGSPAKDFALLTQ 300
Qy 282 CNHTMTGTGFWAAYLTGGDTIYLANFTLPDSPLKIFKPEAAFLPWTGIAADLSPL 341
Db 301 CNHTMTGTGFWAAYLAGGDTIYLANFTLPDSPLKIFKPEAAFLPWTGIAADLSPL 360

RESULT 9
FUT1_HUMAN STANDARD; PRT; 365 AA.
ID FUT1_HUMAN
AC P19526;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1) (Blood group H alpha 2-fucosyltransferase).
GN FUT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370848; PubMed=2118655;
RA Larsen R.D., Ernst L.K., Nair R.P., Lowe J.B.;
RT "Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can

```



```

form the H blood group antigen."
Proc. Natl. Acad. Sci. U.S.A. 87:6674-6678(1990).
[2]
SEQUENCE FROM N.A., AND VARIANT ALLELES.
RP MEDLINE=97240210; PubMed=9122901;
RA Wagner F.F., Flegel W.A.;
RT "Polymorphism of the h allele and the population frequency of
RT sporadic nonfunctional alleles."
RL Transfusion 37:284-290(1997).
[3]
VARIANT HIS-164.
RP MEDLINE=94286534; PubMed=7912436;
RA Kelly R.J., Ernst L.K., Larsen R.D., Bryant J.G., Robinson J.S.,
RA Lowe J.B.;
RT "Molecular basis for H blood group deficiency in Bombay (Oh) and
RT para-Bombay individuals."
RL Proc. Natl. Acad. Sci. U.S.A. 91:5843-5847(1994).
[4]
VARIANT ARG-242.
RP MEDLINE=97445117; PubMed=9299444;
RA Koda Y., Soejima M., Johnson P.H., Smart E., Kimura H.;
RT "Missense mutation of FUT1 and deletion of FUT2 are responsible for
RT Indian Bombay phenotype at ABO blood group system."
RL Biochem. Biophys. Res. Commun. 238:21-25(1997).
CC -!- FUNCTION: CREATES A SOLUBLE PRECURSOR OLIGOSACCHARIDE FUC-ALPHA
CC ((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN ESSENTIAL
CC SUBSTRATE FOR THE FINAL STEP IN THE SOLUBLE A AND B ANTIGEN
CC SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE SAME ACCEPTOR
CC SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
CC alpha-L-fucosyl-1,2-beta-D-galactosyl-R.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- POLYMORPHISM: NONFUNCTIONAL MUTANT OF FUT1 ARE THE CAUSE OF THE H-
CC BOMBAY BLOOD GROUP.
CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
CC TISSUE-SPECIFIC MANNER WITH EXPRESSION RESTRICTED TO CELLS OF
CC MESODERMAL OR ENDODERMAL ORIGIN RESPECTIVELY.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M35531; AAA52639.1; -.
CC EMBL; Z69587; CAA93435.1; -.
CC PIR; A36047; A36047.
CC Genew; HGNC:4012; FUT1.
CC MIM; 211100; -.
CC Interpro; IPR002516; GT_11.
CC Pfam; PF01531; GlycoTransf_11; 1.
CC Kunitz; K00001; Glycosyltransferase; Glycoprotein; Transmembrane;
CC Transferase; Glycosyltransferase; Polymorphism; Blood group antigen.
CC Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.
CC DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 9 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC DOMAIN 26 365 LUMENAL, CATALYTIC (POTENTIAL).
CC DOMAIN 65 65 CARBOHYD (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC VARIANT 154 154 Y -> C (IN BOMBAY H-).
CC /FtId=VAR_003417.
CC L -> H (IN PARA-BOMBAY).
CC /FtId=VAR_009708.
CC W -> C (IN BOMBAY H-).
CC /FtId=VAR_003418.
CC L -> R (IN BOMBAY H-).
CC /FtId=VAR_009709.
CC V -> E (IN BOMBAY H-).

```

```

FT VARIANT 315 315 /FtId=VAR_003419.
FT A -> V (IN BOMBAY H-).
FT /FtId=VAR_003420.
FT W -> C (IN BOMBAY H-).
FT /FtId=VAR_003421.
SQ SEQUENCE 365 AA; 41251 MW; 4F4442BC375C9D9E CRC64;
Query Match 58.4%; Score 1080.5; DB 1; Length 365;
Best Local Similarity 59.7%; Pred. No. 2.8e-85;
Matches 209; Conservative 41; Mismatches 79; Indels 21; Gaps 4;
QY 12 MAHFILFVTVSTIFHVQQR-----LAKIQAMWELPVOIPVL-----ASTSK 53
DB 11 LAFLLVCVLSVIFFLIHQDSIFPHGLGLSILCPDRRLVTPPPVAIFCLPGTAMGPNASSSC 70
QY 54 ALGPSOLRGMTTINAIGRNGNMGEXATYALAKMNGRPAFIPQAOMHSTLAPIFRTITLV 113
DB 71 PQHPASLSGTWTVVPGNGRFGNQMGQVATLLAQLNGRRRAFILPAMHAALAPVFRITLV 130
QY 114 LHSATASRTPQWYHNLNDWMEYRHPGCEYVRFYPCSWTFYHHLRQELI-CEFTLHD 173
DB 131 LAPEVDSRTWPRELQLHDWMESEYADL-RDPFLKLSGFCPSWTFHHLRQELIREFTLHD 189
QY 174 HVREEAQKELRGLOV--NGSRPCTFVGVHVRGDIYVHMPKVKVGVVADRRYILQCALDWF 231
DB 190 HLREEAQSVLQGLRGTRGTRPRFVGVHVRGDIYVHMPKVKVGVVADRRYILQCALDWF 249
QY 232 RARYSSLIFFVTSNGMAWCRENIDTSHGVDYVFGDGIEGSPAKDFALLTQCNTHTMTIGT 291
DB 250 RARHEAPVFFVTSNGMWCENIDTSGDVTFAAGDQGEATPWKDFALLTQCNTHTMTIGT 309
QY 292 FGTHAAVLTGGDIYIYANYILPDSFPLKIFKPAFAFLPEWTGTAADLSPL 341
DB 310 FGFWAAVLAGDGVILANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL 359
RESULT 10
FUT1_PIG STANDARD; PRT; 365 AA.
AC Q29043; O19101;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 1 (EC 2.4.1.69) (GDP-L-fucose:beta-
DE D-galactoside 2-alpha-L-fucosyltransferase 1) (Alpha(1,2)FT 1)
DE (Fucosyltransferase 1).
DE FUT1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96217559; PubMed=8613146;
RA Cohnen S., Voughtouris E., McKenzie I.F., Sandrin M.S.;
RT "Molecular cloning of the gene coding for pig alpha1-->2
RL fucosyltransferase."
RN Immunogenetics 44:76-79(1996).
RP SEQUENCE FROM N.A., AND VARIANTS THR-103 AND GLN-286.
RX MEDLINE=97468270; PubMed=9321466;
RA Meijerink E., Fries R., Voegel P., Masabanda J., Wigger G.,
RA Stricker C., Neuenchwander S., Bertschinger H.O., Stranzinger G.;
RT "Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11
RT are closely linked to the blood group inhibitor (S) and Escherichia
RT coli F18 receptor (ECF18R) loci."
RL Mamm. Genome 8:736-741(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Meijerink E., Neuenchwander S., Fries R., Dinter A.,
RA Bertschinger H.U., Stranzinger G., Voegel P.;
RT "Expression and activity of porcine alpha(1,2)fucosyltransferases

```

determine erythrocyte antigen precursor O status and susceptibility to
 determine erythrocyte antigen precursor O status and susceptibility to
Escherichia coli F18 colonization in the small intestine.;
 Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + beta-D-galactosyl-R = GDP +
 CC alpha-L-fucosyl-1,2-beta-D-galactosyl-1-R.
 CC -!- PATHWAY: Glycosylation
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI.
 CC -!- MISCELLANEOUS: THERE ARE TWO GENES (FUT1 AND FUT2) WHICH ENCODE
 CC GALACTOSIDE 2-L-FUCOSYLTRANSFERASE. THEY ARE EXPRESSED IN A
 CC TISSUE-SPECIFIC MANNER.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 11.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L50534; AAB02984.1; -;
 DR EMBL; U70883; AAB1884.1; -;
 DR EMBL; AF136896; AAF59833.1; -;
 DR InterPro; IPR002516; GT-11.
 DR Pfam; PF01531; Glyco.transf.11; 1.
 DR
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack; Polymorphism.
 KW DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 9 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT FT DOMAIN 29 365 LUMENAL, CATALYTIC (POTENTIAL).
 FT FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT VARIANT 103 103 A -> T.
 FT VARIANT 286 286 R -> Q.
 FT FT CONFLICT 77 77 L -> F (IN REF. 1).
 FT FT CONFLICT 365 AA; 41106 MW; DAFCE77E89A29D75 CRC64;
 SQ SEQUENCE

Db	283	ENTPODDLAYSHYSCDAILITAPSTFGWNLGYLSKGOAVYYQDIRSTNDNYKKGVLPD	342
Qy	324	EAAFLPEWTGIAAD	337
Db	343	DFFVPSWTSIMLD	356
 RESULT 13 FUT8_PIG ID FUT8_PIG STANDARD; PRT; 575 AA. AC P79282; DT 15-JUN-2002 (Rel. 41, Created) DT 15-JUN-2002 (Rel. 41, Last sequence update) DE Alpha-(1,6)-fucosyltransferase (EC 2.4.1.68) (glycoprotein 6-alpha-L-fucosyltransferase) DE (GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1,6-fucosyltransferase) DE (alpha1-6Fuci) (Fucosyltransferase 8). OS Sus scrofa (Pig). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. OX NCBI_TaxID=9823; RN [1] RP SEQUENCE FROM N.A., SEQUENCE OF 102-130; 333-344; 402-415 AND 566-575, AND FUNCTION. RC TISSUE=Brain; RX MEDLINE=97066976; PubMed=8910378; RA Uozumi N., Yanagidani S., Miyoshi E., Ihara Y., Sakuma T., Gao C.-X., Teshima T., Fujii S., Shiba T., Taniguchi N.; RT "Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1->6fucosyltransferase"; RL J. Biol. Chem. 271:27810-27817(1996). CC -!- FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to the first GlcNAc residue, next to the peptide chains in N-glycans. The pH optimum is 7. CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-glucosaminyl-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-beta-N-acetyl-1,4-D-glucosaminyl-1,4-N-acetyl-D-glucosaminyl)asparagine = GDP + N4-(N-acetyl-beta-D-glucosaminyl-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-N-acetyl-D-glucosaminyl)asparagine. CC -!- PATHWAY: Glycosylation. CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound form in trans cisternae of Golgi (By similarity). CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN. CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN. CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23. This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). EMBL; D86723; BAA13157.1; . InterPro: IPR001452; SH3. Pfam: PF00018; SH3; 1. SMART: SM00326; SH3; 1. PROSITE: PS00002; SH3; 1. Transferrase; Glycosyltransferase; Transmembrane; Signal-anchor; Golgi stack; SH3 domain. DOMAIN 1 9 TRANSMEM 10 30 FT DOMAIN 31 575 FT DOMAIN 502 563 FT SITE 299 303 FT SITE 365 366 FT FT FT			

```
SQ SEQUENCE 575 AA; 66229 MW; 0F199D0BC2018F7B CRC64;
Query Match 5.08; Score 92; DB 1; Length 575;
Best Local Similarity 20.5%; Pred. No. 2.6;
Matches 60; Conservative 47; Mismatches 108; Indels 78; Gaps 16;
QY 14 HFLFVFTVSVIFHVOQLAKIQAWEL-----PVQIPVLASTSKALGPSQLRGMWTIN 67
DB 225 HHVVCFMIA--YGTQRLALESHUNRYATGWETVFPVSETCTDRSGSS--TGHWS-- 278
QY 68 AIGRLGNOMGEYATLYALAKNNGRPAFTPAQMHSTLAFIRITPLVLSHATRASRIPWQNY 127
DB 279 --GEVKDKNVQVVELPIVDSVHPRPYLPPLAVPEDLAD-----RLRVHGDPA--VWVWSQ 330
QY 128 HL-----NDWMESE-----YRHTPPGEYVREFTGPGCSWFYFHHLRQLLEQFTLH 172
DB 331 FVKYLIRPOPMPLEKEIEBATKKGFKHPIGVHVRVTRDKVGAEAFAHP-----IEEYIV- 384
QY 173 DHVREAEQKFLRGLQVNGSR-----PGTFVGVIHVRGDYVHVMPK--VW----- 214
DB 365 -HVEDFQLLARRMOVDKRYVLAITDDPALLKEAKTKYPSYEFISDINSWSAGLHNRYT 443
QY 215 -----KGVVADRYLQOALDWFARYSSLIFFVTSNGMAWCR---ENIDTSHGD 260
DB 444 ENSLRGVILDIHFLSQA-DFLVCTFSQV-----CRVAYEIMOALHPD 485
RESULT 14
ID FUT8_HUMAN STANDARD; PRT; 575 AA.
AC Q9BYC5; Q9BYC6; Q9P2U5; Q9P2U6; Q00235;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-(1,6)-fucosyltransferase (EC 2.4.1.68) (Glycoprotein 6-alpha-L-
DE fucosyltransferase) (GDP-fucose-glycoprotein fucosyltransferase)
DE (GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1,6-fucosyltransferase)
DE (alpha1-6FucT) (Fucosyltransferase 8).
GN FUT8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RX MEDLINE=97279058; PubMed=9133635;
RA Yanagidani S., Uozumi N., Ihara Y., Miyoshi E., Yamaguchi N.,
RA Taniguchi N.;
RT "Purification and cDNA cloning of GDP-L-Fuc:N-acetyl-beta-D-
RT glucosaminide:alpha1-6 fucosyltransferase (alpha1-6 FucT) from human
RT gastric cancer MKN45 cells.";
RN J. Biochem. 121:626-632(1997).
RL [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryo;
RA Caillaud A., Balanzino L., Candellier J.J., Oriol R., Mollicone R.;
RT "Differential splice variants of human FUT8 embryonic cDNA.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=20275614; PubMed=10814706;
RA Yamaguchi Y., Ikeda Y., Takahashi T., Ihara H., Tanaka T., Sasho C.,
RA Uozumi N., Yanagidani S., Inoue S., Fujii J., Taniguchi N.;
RT "Genomic structure and promoter analysis of the human
RT alpha1,6-fucosyltransferase gene (FUT8).";
RL Glycobiology 10:637-643(2000).
RN [4]
RP DONOR SUBSTRATE BINDING, AND MUTAGENESIS OF ARG-365 AND ARG-366.
RX MEDLINE=20229550; PubMed=10764839;
RA Takahashi T., Ikeda Y., Tateishi A., Yamaguchi Y., Ishikawa M.,
RA Taniguchi N.;
RT "A sequence motif involved in the donor substrate binding by
RT alpha1,6-fucosyltransferase: the role of the conserved arginine
residues.";
RL Glycobiology 10:503-510(2000).
-!- FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to
the first GlcNAc residue, next to the peptide chains in N-glycans.
-!- CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-
glucosaminyl-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-
mannosyl-beta-N-asparaginyl-1,4-D-glucosaminyl-1,4-N-acetyl-D-
glucosaminyl)asparagine = GDP + N4-(N-acetyl-beta-D-glucosaminyl-
1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-
beta-N-acetyl-D-glucosaminyl-1,4-(alpha-L-fucosyl-1,6)-N-acetyl-D-
glucosaminyl)asparagine.
-!- PATHWAY: Glycosylation.
-!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
form in trans cisternae of Golgi (By similarity).
-!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/retinal;
are produced by alternative splicing.
-!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; D89289; BAA19764.1; -
EMBL; AF038281; AB92372.2; -
EMBL; AF038280; AB92372.2; JOINED.
EMBL; Y17975; CAA76988.1; -
EMBL; Y17976; CAA76985.1; -
EMBL; Y17977; CAA76986.1; -
EMBL; Y17978; CAA76987.1; -
EMBL; AB049828; BAB40975.1; -
EMBL; AB049740; BAB40929.2; -
EMBL; AB032573; BAA92859.2; -
EMBL; AB032572; BAA92858.1; -
EMBL; AB032568; BAA92858.1; JOINED.
EMBL; AB032569; BAA92858.1; JOINED.
EMBL; AB032570; BAA92858.1; JOINED.
EMBL; AB032571; BAA92858.1; JOINED.
EMBL; HGNC:4019; FUT8.
MIM; 602589; -
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; FALSE_NEG.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
KW Golgi_stack; SH3 domain; Alternative splicing.
FT DOMAIN 1 9
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 10 30
FT (POTENTIAL).
FT LUMENAL, CATALYTIC (POTENTIAL).
FT DOMAIN 31 575
FT SH3-BINDING (POTENTIAL).
FT SITE 299 305
FT SITE 365 366
FT VARSPLIC 280 329
FT E (IN ISOFORM 2).
FT VARSPLIC 330 575
FT MUTAGEN 365 365
FT MUTAGEN 366 366
SQ SEQUENCE 575 AA; 66515 MW; 5AE24A93881E18D0 CRC64;
Query Match 4.9%; Score 90; DB 1; Length 575;
Best Local Similarity 20.2%; Pred. No. 3.8;
Matches 59; Conservative 46; Mismatches 111; Indels 76; Gaps 15;
QY 14 HFLFVFTVSV-----TIFHVQQLAKIQAWELFVQIPVLASTSKALGPSQLRGMWTINA 68
DB 225 HHVVCFMIA-YGTQRLALESHUNRYATGWETVFPVSETCTDRSGIS--TGHWS--- 278
```

```
QY 69 IGRGNQMGVATYALAKMNGRPAFIPAQMHSHTLAPIERTLPLVLSATASRIPQWNYH 128
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 -GEVADKNVQVVELPIVDSLHPRPPIYPLAVPEDLAD-----RLVVRHGDPD--VWVWSQF 331
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 L-----NDWMEEE-----YRHPPGIVRFTGYPCSWTFYHHLRQBILOEFTLHD 173
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 332 VKYLTPQWLEKEIEEATKKGFKHPVIGVHVVRTDKVGTEAAFPPIEEYMV----- 384
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 HVREBAQKFLRGLQVNGSR-----PGTFVGVHVRGDIYVHVMPK--VW----- 214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 HVEEHFQLLARRMQVDKKRVYLATDDPSLLKEATKYNYEFISDNTSWSAGLHNRYTE 444
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 --KGVDADRYLQALDWFRRARYSSLLFVVTSCMAWCR---ENIDTSHGD 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 445 NSLRGVILDIHFLSQD-DFLVCTFSSQV-----CRVAYEIMQTLHPD 485
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
FUT8_MOUSE
ID FUT8_MOUSE STANDARD; PRT; 575 AA.
AC Q9WIS2; Q921U1;
DT 15-JUN-2002 (Rel. 41, Crested)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-(1,6)-fucosyltransferase (EC 2.4.1.68) (Glycoprotein 6-alpha-L-
fucosyltransferase) (GDP-fucose-glycoprotein fucosyltransferase)
DE (GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1,6-fucosyltransferase)
DE (alpha1-6Fuc) (Fucosyltransferase 8).
CN FUT8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20358720; PubMed=10902914;
RA Hayashi H., Yoneda A., Asada M., Ikekita M., Imamura T.;
RT "Molecular cloning of mouse alpha-1,6-fucosyltransferase and
expression of its mRNA in the developing cerebrum.";
RL DNA Seq. 11:91-96(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the addition of fucose in alpha 1-6 linkage to
the first GlcNAc residue, next to the peptide chains in N-glycans
(CC By similarity).
CC -!- CATALYTIC ACTIVITY: GDP-L-fucose + N4-(N-acetyl-beta-D-
glucosaminyl-1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-
mannosyl-beta-N-acetyl-1,4-D-glucosaminyl-1,4-N-acetyl-D-
glucosaminyl)asparagine -> GDP + N4-(N-acetyl-beta-D-glucosaminyl-
1,2-alpha-D-mannosyl-1,3-(R(1)-alpha-1,6)-beta-D-mannosyl-1,4-
beta-N-acetyl-D-glucosaminyl-1,4-(alpha-L-fucosyl-1,6)-N-acetyl-D-
glucosaminyl)asparagine.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
form in trans cisternae of Golgi (By similarity).
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 23.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB025198; BAA76392.1; -
DR EMBL; BC010666; AAH10666.1; -
DR MGD; MGI:1858901; Fut8.
DR InterPro; IPR001452; SH3.
```

```
DR Pfam: PF00018; SH3; 1.
DR SMART: SMO0326; SH3; 1.
DR PROSITE; PS00000; SH3; FALSE_NEG.
KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
FT Golgi stack; SH3 domain.
FT DOMAIN 1 9
FT TRANSMEM 10 30
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ET DOMAIN 31 575
ET DOMAIN 502 563
FT SITE 299 305
FT SITE 365 366
FT SITE 40 40
FT CONFLICT 388 388
FT CONFLICT 418 418
FT CONFLICT N -> K (IN REF. 2).
FT CONFLICT N -> E (IN REF. 2).
SQ SEQUENCE 575 AA; 66555 MW; 7BE2ED3146E0C45F CRC64;

Query Match 4.9%; Score 90; DB 1; Length 575;
Best Local Similarity 19.8%; Pred. No. 3.8;
Matches 58; Conservative 49; Mismatches 108; Indels 78; Gaps 16;

QY 14 HFILEVFTVS-----TIFHVQORLAKIQMWELPVOIPVLASTSKALGPSQLRGWWTINA 68
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 HHVVYCFMIAYGTQRTLILESQNWRYATGGWE-TVFRPVSETCTDRSGLS--TGHWS--- 278
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 IGRGNQMGVATYALAKMNGRPAFIPAQMHSHTLAP-IFRITLPLVLSATASRIPQWNY 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 -GEVADKNVQVVELPIVDSLHPRPPIYPLAVPEDLADRLRV-----HGDPD--VWVWSQ 330
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 128 HL-----NDWMEEE-----YRHPPGIVRFTGYPCSWTFYHHLRQBILOEFTLH 172
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 331 FVKYLTPQWLEKEIEEATKKGFKHPVIGVHVVRTDKVGTEAAFPPIEEYMV----- 384
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 DHVREBAQKFLRGLQVNGSR-----PGTFVGVHVRGDIYVHVMPK--VW----- 214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 -HVEEHFQLLARRMQVDKKRVYLATDDPTLLKEATKYNYEFISDNTSWSAGLHNRYT 443
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 ---KGVDADRYLQALDWFRRARYSSLLFVVTSCMAWCR---ENIDTSHGD 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 444 ENSLRGVILDIHFLSQD-DFLVCTFSSQV-----CRVAYEIMQTLHPD 485
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 27, 2003, 15:07:17
Job time : 13.0683 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:01:52 ; Search time 36.6774 Seconds
(without alignments)
1932.533 Million cell updates/sec

Title: US-10-040-863-11

Perfect score: 1850

Sequence: 1 MLVQMPSFPMHILFVF.....AAFLPEWTGTAADLSPLLK 344

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1825.5	98.7	343	6	Q9TUD3
2	1821.5	98.5	343	6	Q77486
3	1815.5	98.1	343	6	Q77485
4	1793.5	96.9	338	6	Q9TUD5
5	1784.5	96.5	343	6	Q9TTC7
6	1771.5	95.8	332	4	Q99450
7	1766.5	95.5	343	6	Q77487
8	1736.5	93.9	343	6	Q77712
9	1566.5	84.7	346	4	Q14338
10	1521.5	82.2	347	6	Q29505
11	1483	80.2	354	11	Q35087
12	1481	80.1	354	11	Q9R275
13	1475.5	79.8	347	11	Q9JL27
14	1475	79.7	380	11	Q9JK44
15	1471.5	79.5	347	11	Q70504
16	1409.5	76.2	273	6	Q9TSL5

17	1376	74.4	353	11	O55025
18	1369.5	74.0	321	11	Q91VF0
19	1368.5	74.0	321	11	Q920W2
20	1354.5	73.2	321	11	Q920W3
21	1344.5	72.7	273	6	Q9TTC9
22	1335.5	72.2	273	6	Q9TUD0
23	1332.5	72.0	348	6	Q9TUD1
24	1311.5	70.9	332	6	Q9N266
25	1304.5	70.5	368	11	Q920T3
26	1301.5	70.4	368	11	Q920T4
27	1298.5	70.2	368	11	Q91V73
28	1295.5	70.0	368	11	Q920T2
29	1293.5	69.9	368	6	Q9TTC3
30	1233.5	66.7	299	11	Q9WUE6
31	1100.5	59.5	360	6	Q9TTC7
32	1092	59.0	377	11	Q9JL28
33	1092	59.0	377	11	Q9T327
34	1084	58.6	366	6	Q9TUD4
35	1082	58.5	377	11	Q920W5
36	1082	58.4	365	6	Q9TUE7
37	1080.5	58.4	365	6	Q9TUE7
38	1076.5	58.2	365	4	Q8WXB6
39	1076	58.2	366	6	Q9TUE6
40	1075.5	58.1	365	4	O14505
41	1072	57.9	366	6	Q9TUD6
42	1071.5	57.9	365	4	O14506
43	1071.5	57.9	365	4	O14507
44	1071	57.9	366	6	Q9TUD2
45	1070.5	57.9	353	6	Q9N270

ALIGNMENTS

RESULT 1

Q9TUD3	
ID	Q9TUD3
AC	Q9TUD3; PRELIMINARY; PRT; 343 AA.
DT	01-MAY-2000 (TRENBLrel. 13, Created)
DT	01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE	01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE	Alpha (1,2) fucosyl transferase.
GN	FUT2
OS	Gorilla gorilla (gorilla).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
OX	NCBI_TaxID=9593;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-ALEXIS;
RX	MEDLINE-20188794; PubMed=10723735;
RA	Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
RA	Blancher A.;
RT	"Evolution of alpha2-Fucosyltransferase genes in primates: relation
RT	between an intronic Alu-Y element and red cell expression of ABH
RL	antigens.";
RL	Mol. Biol. Evol. 17:337-351(2000).
DR	EMBL; AF080606; AAF14068.1;
DR	InterPro; IPR002516; GT_11.
DR	Pfam; PF01531; Glyco_transf_11; 1.
KW	Transferase.
SK	SEQUENCE 343 AA; 39001 MW; A753375D47AE8C8 CRC64;

Query Match	98.7%; Score 1825.5; DB 6; Length 343;
Best Local Similarity	99.4%; Pred. No. 6.7e-162;
Matches	342; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy	1 MLVQMPSFPMHILFVFVTSTIFHVQQLAKIQAMWELPVQIPVLASTSKALGPSQL 60
Db	1 MLVQMPSFPMHILFVFVTSTIFHVQQLAKIQAMWELPVQIPVLASTSKALGPSQL 60
Qy	6- RGMWTINAIKRLGNMGYATLYALAKNGRPAPFPAQMHSTLAFIFRITLPVLHSATAS 120

```
Db 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
QY 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 180
Db 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSLIIF 240
Db 180 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 344
Db 300 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 343

RESULT 2
O77486 PRELIMINARY; PRT: 343 AA.
AC O77486;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015635; BAA31128.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38973 MW; 894E28BD74AE8FBC CRC64;

Query Match 98.5%; Score 1821.5; DB 6; Length 343;
Best Local Similarity 99.1%; Pred. No. 1.6e-161;
Matches 341; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MLVVQMPFSPMAHFILFVFTVSTIFHVQORLAKIOAMWELPVOIPVLASTSKALGPSQL 60
Db 1 MLVVQMPFSPMAHFILFVFTVSTIFHVQORLAKIOAMWELPVOIPVLASTSKALGPSQL 60
QY 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
Db 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
QY 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 180
Db 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSLIIF 240
Db 180 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299

RESULT 4
Q9TUD5 PRELIMINARY; PRT: 338 AA.
ID Q9TUD5
AC Q9TUD5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
```

```
Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 344
Db 300 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 343

RESULT 3
O77485 PRELIMINARY; PRT: 343 AA.
AC O77485;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Alpha(1,2) fucosyltransferase.
GN FUT2.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Koda Y., Tachida H., Soejima M., Takenaka O., Kimura H.;
RT "The old origin of a null allele se428 of the human ABO-secretor type
RT alpha(1,2) fucosyltransferase gene (FUT2).";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; AB015634; BAA31127.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 343 AA; 38999 MW; B849D7E152852081 CRC64;

Query Match 98.1%; Score 1815.5; DB 6; Length 343;
Best Local Similarity 98.8%; Pred. No. 5.8e-161;
Matches 340; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MLVVQMPFSPMAHFILFVFTVSTIFHVQORLAKIOAMWELPVOIPVLASTSKALGPSQL 60
Db 1 MLVVQMPFSPMAHFILFVFTVSTIFHVQORLAKIOAMWELPVOIPVLASTSKALGPSQL 60
QY 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
Db 61 RGMWTINAIGRLGNQMGGEATLYALAKMNGRPAFIPQAMHSTLAPIFRITLPVLHSATAS 120
QY 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 180
Db 121 RIPQWNYHLNDWMEEXYRHIPPEYVRETYGPCSWTFYHHLRQELQEFTHLHDHVRREAQ 179
QY 181 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSLIIF 240
Db 180 KFLRGLQVNSRPGTFVGVHVRRGDYVHVMPKWKGVVADRYLQOALDWFRRYSSPIF 239
QY 241 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 300
Db 240 VVTSNGMAWCRENIDTSHGDVVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTFGIWAAYLT 299
QY 301 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 344
Db 300 GGDITYLANYTLPDPSFLKIKPEAFLEPWTGIAADLSPLKKH 343

RESULT 4
Q9TUD5 PRELIMINARY; PRT: 338 AA.
ID Q9TUD5
AC Q9TUD5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
```

RT between an intronic Alu-Y element and red cell expression of ABH
 RT antigens.";
 RL Mol. Biol. Evol. 17:337-351(2000).
 DR EMBL: AF136648; AAF25585.1; -.
 DE Alpha (1,2) fucosyl transferase.
 GN FUT2.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH16;
 RX MEDLINE=20188794; PubMed=10723735;
 RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
 BA Blancher A.;
 RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation
 RT between an intronic Alu-Y element and red cell expression of ABH
 RT antigens.";
 RL Mol. Biol. Evol. 17:337-351(2000).
 DR EMBL: AF080604; AAF14066.1; -.
 DE InterPro: IPR002516; GT.11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Transferase.
 SQ SEQUENCE 338 AA; 38428 MW; E7FABA0FF1BC95F9 CRC64;
 Query Match 96.9%; Score 1793.5; DB 6; Length 338;
 Best Local Similarity 98.8%; Pred. No. 6.4e-159;
 Matches 335; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 6 MPSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOLRGMT 65
 Db 1 MPSPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOLRGMT 60
 QY 66 INALGRGNMGVATLYALAKNGRPAFIPAQMHSFLAPFIRITLPLVLSATASRIPWQ 125
 Db 61 INALGRGNMGVATLYALAKNGRPAFIPAQMHSFLAPFIRITLPLVLSATASRIPWQ 120
 QY 126 NYHLNDWMEERYRHIPGEYVRTGPGSWTFYHHLRQEILOEFTLHDHVRREAQKFLRG 185
 Db 121 NYHLNDWMEERYRHIPGEYVRTGPGSWTFYHHLRQEILOEFTLHDHVRREAQKFLRG 179
 QY 186 LOVNGSRPGTFVGVHVRGDIYVHVMPKVGVDVRRYLOQALDWFARYSLLIFVTSN 245
 Db 180 LOVNGSRPGTFVGVHVRGDIYVHVMPKVGVDVRRYLOQALDWFARYSLLIFVTSN 239
 QY 246 GMACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLTGGDTI 305
 Db 240 GMACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLTGGDTI 299
 QY 306 YLANITLTPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 344
 Db 300 YLANITLTPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 338
 RESULT 5
 Q9TTC7 PRELIMINARY; PRT; 343 AA.
 ID Q9TTC7
 AC Q9TTC7
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Alpha(1,2)fucosyltransferase.
 GN FUT2.
 OS Hylobates lar (Common gibbon).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=9580;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TJGA;
 RX MEDLINE=20188794; PubMed=10723735;
 RA Apoll P.-A., Roubinet F., Despiau S., Mollicone R., Oriol R.,
 RA Blancher A.;
 RT "Evolution of alpha2-Fucosyltransferase genes in primates: relation

RT between an intronic Alu-Y element and red cell expression of ABH
 RT antigens.";
 RL Mol. Biol. Evol. 17:337-351(2000).
 DR EMBL: AF136648; AAF25585.1; -.
 DE InterPro: IPR002516; GT.11.
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 343 AA; 39059 MW; A52161BE67E1F64D CRC64;
 Query Match 96.5%; Score 1784.5; DB 6; Length 343;
 Best Local Similarity 97.4%; Pred. No. 4.5e-158;
 Matches 335; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
 QY 1 MLVVQMPFPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60
 Db 1 MLVVQMPFPFMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOL 60
 QY 61 RGMWTINAGRLGNMGVATLYALAKNGRPAFIPAQMHSFLAPFIRITLPLVLSATAS 120
 Db 61 RGMWTINAGRLGNMGVATLYALAKNGRPAFIPAQMHSFLAPFIRITLPLVLSATAS 120
 QY 121 RIPWQNYHLNDWMEERYRHIPGEYVRTGPGSWTFYHHLRQEILOEFTLHDHVRREAQ 180
 Db 121 RIPWQNYHLNDWMEERYRHIPGEYVRTGPGSWTFYHHLRQEILOEFTLHDHVRREAQ 179
 QY 181 KFLRGLOVNGSRPGTFVGVHVRGDIYVHVMPKVGVDVRRYLOQALDWFARYSLLIF 240
 Db 180 KFLRGLOVNGSRPGTFVGVHVRGDIYVHVMPKVGVDVRRYLOQALDWFARYSLLIF 239
 QY 241 VVTSNGAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLT 300
 Db 240 VVTSNGAWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNTHTIMTIGTGIWAAYLT 299
 QY 301 GGDFTYLANITLTPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 344
 Db 300 GGDFTYLANITLTPDSPFLKIFKPEAAFLPEWTGIAADLSPLLK 343
 RESULT 6
 Q99450 PRELIMINARY; PRT; 332 AA.
 ID Q99450
 AC Q99450;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Alpha (1,2) fucosyltransferase.
 GN SEJ.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96199252; PubMed=8621666;
 RA Kudo T., Iwasaki H., Shinya N., Nishihara S., Ando T., Narimatsu I.,
 RA Narimatsu H.;
 RT "Molecular Genetic Analysis of Human Histo-blood Group System. II.
 RT Secretor Gene Inactivation by a Novel Single Missense Mutation A365T
 RT in Japanese Nonsecretor Individuals.";
 RL J. Biol. Chem. 271:9830-9837(1996).
 DR EMBL: D89326; BAA13943.1; -.
 DR InterPro: IPR002516; GT.11
 DR Pfam: PF01531; Glyco_transf_11; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 332 AA; 37773 MW; 092A4C317DA45986 CRC64;
 Query Match 95.8%; Score 1771.5; DB 4; Length 332;
 Best Local Similarity 99.4%; Pred. No. 7e-157;
 Matches 331; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 12 MAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOLRGMTNAIGR 71
 Db 1 MAHFILFVFTVSTIFHVQORLAKIQAMWELPVQIPVLASTSKALGPSOLRGMTNAIGR 60


```
Q14338
ID Q14338 PRELIMINARY; PRT; 346 AA.
AC Q14338;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha (1, 2) fucosyltransferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP PARTIAL SEQUENCE FROM N.A.
RA Kimura H.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE OF 84-106 FROM N.A.
RX MEDLINE=96335703; PubMed=875920;
RA Koda Y., Soejima M., Liu Y., Kimura H.;
RT "Molecular basis for secretor type alpha(1,2)-fucosyltransferase gene
RT deficiency in a Japanese population: a fusion gene generated by
RT unequal crossover responsible for the enzyme deficiency.";
RL Am. J. Hum. Genet. 59:343-350(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181460; PubMed=7876235;
RA Kelly R.J., Rouquier S., Giorgi D., Lennon G.G., Lowe J.B.;
RT "Sequence and expression of a candidate for the human secretor blood
RT group alpha(1,2) fucosyltransferase gene (FUT2).";
RL J. Biol. Chem. 270:4640-4649(1995).
DR EMBL; D82933; BA011638.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferrase.
SQ SEQUENCE 346 AA; 39395 MW; 9D72275F8571410B CRC64;

Query Match 84.7%; Score 1566.5; DB 4; Length 346;
Best Local Similarity 87.3%; Pred. No. 9.7e-138;
Matches 295; Conservative 9; Mismatches 23; Indels 11; Gaps 3;

QY 9 SFPKAFILFVFFVSTIFHVOORLAKTOAMWELPVQIPVLASTKALGPSOL--RGMWTI 66
DB 18 SFTFFYFAIFVFFVSTIFHVOORLAKTOAMWELPVQIPVLASTKALGPSOL--RGMWTI 69

QY 67 NAIKRLGNQMGVATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLSATSRIPWQN 126
DB 70 NSKRLGNQMGVATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLSATSRIPWQN 129

QY 127 YHLNDWMEERYRHIPPGVYRFTGYPGCSWTFFYHLRQELIQEFTLHDHVRREAAKFLRGL 186
DB 130 YHLNDWMEERYRHIPPGVYRFTGYPGCSWTFFYHLRQELIQEFTLHDHVRREAAKFLRGL 188

QY 187 QVNSRPGCTFVGHVHRGDYVHVMPKWKGVVADRRYLQALDWFRRYSLLIFVWTSNG 246
DB 189 QVNSRPGCTFVGHVHRGDYVHVMPKWKGVVADRRYLQALDWFRRYSLLIFVWTSNG 248

QY 247 NAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFFGIWAAYLTGGDTIY 306
DB 249 NAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFFGIWAAYLTGGDTIY 308

QY 307 LANTLPDSPFLKIFKPEAFLPWTGTAADLSPLKKH 344
DB 309 LANTLPDSPFLKIFKPEAFLPWTGTAADLSPLKKH 346

RESULT 10
Q29505
ID Q29505 PRELIMINARY; PRT; 347 AA.
AC Q29505;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Galactoside 2-L-fucosyltransferase 3 (EC 2.4.1.69) (Secretor blood
```

```
DE group ALPHA-2-fucosyltransferase) (GDP-L-fucose:beta-D-galactoside
DE 2-ALPHA-L-fucosyltransferase) (ALPHA(1,2)F) (Fucosyltransferase 3).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=96279281; PubMed=8663168;
RA Hitoshi S., Kojima N., Kanazawa I., Tsuji S.;
RT "Molecular cloning and expression of a third type of rabbit GDP-L-
RT fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase.";
RL J. Biol. Chem. 271:16975-16981(1996).
CC -!- FUNCTION: CREATES A MEMBRANE-ASSOCIATED PRECURSOR OLIGOSACCHARIDE
CC FUC-ALPHA((1,2)GALBETA-) CALLED THE H ANTIGEN WHICH IS AN
CC ESSENTIAL SUBSTRATE FOR THE FINAL STEP IN THE MEMBRANE-ASSOCIATED
CC A AND B ANTIGEN SYNTHESIS PATHWAY. H AND SE ENZYMES FUCOSYLATE THE
CC SAME ACCEPTOR SUBSTRATES BUT EXHIBIT DIFFERENT KM VALUES.
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + BETA-D-GALACTOSYL-R = GDP +
CC ALPHA-L-FUCOSYL-1,2-BETA-D-GALACTOSYL-R.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- TISSUE SPECIFICITY: ADULT GASTROINTESTINAL TRACT AND COLON.
CC -!- MISCELLANEOUS: THERE ARE THREE GENES (FUT1, FUT2 AND FUT3) WHICH
CC ENCODE GALACTOSIDE 2-L-FUCOSYLTRANSFERASE IN RABBIT. THEY ARE
CC EXPRESSED IN A TISSUE-SPECIFIC MANNER.
CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
DR EMBL; X91269; CA062669.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 29 347
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 347 AA; 39469 MW; 746F7007309862A5 CRC64;

Query Match 82.2%; Score 1521.5; DB 6; Length 347;
Best Local Similarity 81.6%; Pred. No. 1.9e-133;
Matches 284; Conservative 19; Mismatches 40; Indels 5; Gaps 3;

QY 1 MLVYOMPFSRPMHFIILFVFTVSTIFHVOORLAKTOAMWE--LPVQIPVLA--STSKALG 56
DB 1 MSTAQVFPFAPFWHVILFVFTASTIFHLCQRLVRIQPTWELLPALIPAVTFRPTSQAP 60

QY 57 PSQLRGWTINAIKRLGNQMGVATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLSH 116
DB 61 SRPLGGMTINAMRLGNQMGVATLYALAKMNGRPAPFPAQMHSHTLAPIFRITLPVLSH 120

QY 117 ATASRIPWNTLNDWMEERYRHIPPGVYRFTGYPGCSWTFFYHLRQELIQEFTLHDHVR 176
DB 121 STASRVPWNTLNDWMEERYRHIPPGVYRFTGYPGCSWTFFYHLRQELIQEFTLHDHVR 179

QY 177 EEAQFLRGLGVNSRPGCTFVGHVHRGDYVHVMPKWKGVVADRRYLQALDWFRRYS 236
DB 180 EEAQFLRGLGVNSRPGCTFVGHVHRGDYVHVMPKWKGVVADRRYLQALDWFRRYS 239

QY 237 SLIFVVTSGNMAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFFGIWA 296
DB 240 SPVFFVVTSGNMAWCRENIDTSHGDVVFAGDGIKSPAKDFALLTQCNTIMTIGTFFGIWA 299

QY 297 AYLTGGDTIYIANYTLPDSPFLKIFKPEAFLPWTGTAADLSPLKKH 344
DB 300 AYLTGGDTIYIANYTLPDSPFLKIFKPEAFLPWTGTAADLSPLKKH 347
```



```
RESULT 11
Q35087 PRELIMINARY; PRT; 354 AA.
AC O35087;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON CANCER;
RA Sejima M., Wang B., Koda Y., Kimura H.;
RT "Two distinct rat GDP-L-fucose:b-D-galactoside 2-a-L-
RL fucosyltransferase genes.";
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006138; BRA21742.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39995 MW; 86364444889215BA1 CRC64;

Query Match 80.2%; Score 1493; DB 11; Length 354;
Best Local Similarity 77.7%; Pred. No. 6.1e-130;
Matches 276; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
DB 1 MASAQVPFSEPLAHFLFVFTVSTIHLQORIVKLOPLSEKELPMTQMSSGNTESP6R 60
QY 50 STSKALGPSOLRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 109
DB 61 RDSEQHNGELRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPASMHNLAFIRI 120
QY 110 TLPVLSATASRIPWQNYHLNDWMEERYHIPGVEYVRFYPCSWTFYHHLRQELIOEF 169
DB 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHI-PGHFVRFYPCSWTFYHHLRPEILKEF 179
QY 170 TLHDHVEEAQKFLRGLOVNGSRPGTFVGVHVRGQYVHVPKWKGVADRYLQOALD 229
DB 180 TLHDHVEEAQKFLRGLOVNGSRPGTFVGVHVRGQYVHVPKWKGVADRYLQOALD 239
QY 230 WFRARYSSLFVYVTSNGMWCNENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTMTI 289
DB 240 MFRARYSSPVFVYVTSNGMWCNENINASRGDVVFAGDIEGSPAKDFALLTQCNTHTMTI 299
QY 290 GTFGIWAAYLTGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWTGIAADLSPLK 344
DB 300 GTFGIWAAYLAGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWVGIPADLSPLK 354

RESULT 12
Q9R275 PRELIMINARY; PRT; 354 AA.
AC Q9R275;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha 1,2-fucosyltransferase B.
GN FTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RA Bureau V., Marionneau S., Cailleteau-Thomas A., Le Moullac-Vaidye B.,
```

```
RA Liehr T., Le Pendu J.;
RT "Comparison of the three rat GDP-L-fucose: beta-D-galactoside 2-
RL alpha-L-fucosyltransferases FTA, FTB and FTC.";
RL Eur. J. Biochem. 268:1006-1019(2001).
DR EMBL; AF131238; AAD24469.1; -.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 354 AA; 39983 MW; 123E8C8379E8559E CRC64;

Query Match 80.1%; Score 1491; DB 11; Length 354;
Best Local Similarity 77.7%; Pred. No. 9.4e-130;
Matches 276; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAM--WELPV-----QIPVLA 49
DB 1 MASAQVPFSEPLAHFLFVFTVSTIHLQORIVKLOPLSEKELPMTQMSSGNTESP6R 60
QY 50 STSKALGPSOLRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPAQMHSILAFIRI 109
DB 61 RDSEQHNGELRGWMTINAGRLGNQMGVEYATLYALAKMNGRPAFIPASMHNLAFIRI 120
QY 110 TLPVLSATASRIPWQNYHLNDWMEERYHIPGVEYVRFYPCSWTFYHHLRQELIOEF 169
DB 121 SLPLVHSDTAKKIPWQNYHLNDWMEERYRHI-PGHFVRFYPCSWTFYHHLRPEILKEF 179
QY 170 TLHDHVEEAQKFLRGLOVNGSRPGTFVGVHVRGQYVHVPKWKGVADRYLQOALD 229
DB 180 TLHDHVEEAQKFLRGLOVNGSRPGTFVGVHVRGQYVHVPKWKGVADRYLQOALD 239
QY 230 WFRARYSSLFVYVTSNGMWCNENIDTSHGDDVVFAGDIEGSPAKDFALLTQCNTHTMTI 289
DB 240 MFRARYSSPVFVYVTSNGMWCNENINASRGDVVFAGDIEGSPAKDFALLTQCNTHTMTI 299
QY 290 GTFGIWAAYLTGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWTGIAADLSPLK 344
DB 300 GTFGIWAAYLAGGDTIYLYANTLPDPSFLKIFKPEAAFLPEWVGIPADLSPLK 354

RESULT 13
Q9JL27 PRELIMINARY; PRT; 347 AA.
AC Q9JL27;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Alpha(1,2)fucosyltransferase FUT2 (EC 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX MEDLINE=21316545; PubMed=11323419;
RA Domino S.E., Zhang L., Lowe J.B.;
RT "Molecular Cloning, Genomic Mapping, and Expression of Two Secretor
RT Blood Group alpha (1,2)fucosyltransferase Genes Differentially
RT Regulated in Mouse Uterine Epithelium and Gastrointestinal Tract.";
RL J. Biol. Chem. 276:23748-23756(2001).
DR EMBL; AF214656; AAF45146.1; -.
DR MGD; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39243 MW; DD5F441046CC8E79 CRC64;

Query Match 79.8%; Score 1475.5; DB 11; Length 347;
Best Local Similarity 79.3%; Pred. No. 3e-129;
Matches 276; Conservative 29; Mismatches 38; Indels 5; Gaps 2;

QY 1 MLVQMPFPMAHFLFVFTVSTIFHVQORLAKIQAMWELPVQIPVLSKALGPSQ- 59
```

```

Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVQMSSPNAARDIMQ 60
QY 60 ---LRGWTINAGRLGNQMGVATYALAKMNGRPAFIPAQMHSITLAPFRITLPLVLS 116
Db 61 SAKLOGIFTINSIGRLGNQMGVATYALAKMNGRPAFIPESMHNALAFIRISLPVLS 120
QY 117 ATASRIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 176
Db 121 DTARRIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 179
QY 177 EEAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALDWRFRYS 236
Db 180 EEAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALDWRFRYS 239
QY 237 SLIFVVTNSGMWACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 296
Db 240 SPVFTVTSNGMWCRENINSLGDDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 299
QY 297 AYLTGSDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 344
Db 300 AYLAGGDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 347

RESULT 14
Q9JK44 PRELIMINARY; PRT; 380 AA.
ID AC Q9JK44;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Alpha 1-2 fucosyltransferase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FTSCHER;
RA Sherwood A.L., Stroud M.R., Levery S.B., Holmes E.H.;
RT "An amino acid region at the N-terminus of rat hepatoma H35 cell GDP-
RT fucose:GMI alpha 1-2 fucosyltransferase is required for optimum enzyme
RT activity and interaction with lipids."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264005; AAF72200.1;
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 380 AA; 42742 MW; 920692C96689C05C CRC64;

Query Match 79.7%; Score 1475; DB 11; Length 380;
Best Local Similarity 77.7%; Pred. No. 3.8e-129;
Matches 275; Conservative 30; Mismatches 37; Indels 12; Gaps 3;

QY 1 MLVQMPSPFMAHILFVFTVSTIIHQRIYVKLQTLSEKELQAVQMSSPNAARDIMQ 49
Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVQMSSPNAARDIMQ 60
QY 50 STSKALGPSQALGWTINAGRLGNQMGVATYALAKMNGRPAFIPAQMHSITLAPFRITLPLVLS 109
Db 61 RDSQHGNGELRGMTFINSIGRLGNQMGVATYALAKMNGRPAFIPESMHNALAFIRISLPVLS 120
QY 110 TLPVLHSATASRIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 169
Db 121 SLPLVLSDFAKKIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 179
QY 170 TLHDHVRQAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALD 229
Db 180 TLHDHVRQAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALD 239
QY 230 WFRARYSSLIFFVTSNGMWCRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 289
Db 240 MFRARYSSPFFVTSNGMWCRENINSLGDDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 299

```

```

QY 290 GTGCIWAAYLTGGDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 343
Db 300 GTGCIWAAYLAGGDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 353

RESULT 15
O70504 PRELIMINARY; PRT; 347 AA.
ID AC O70504;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase (EC
DE 2.4.1.69).
GN FUT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=GASTROINTESTINAL TRACT;
RX MEDLINE=20471982; PubMed=11018479;
RA Lin B., Hayashi Y., Saito M., Sakakibara Y., Yanagisawa M.,
RA Iwamori M.;
RT "GDP-fucose: beta-galactoside alphaL,2-fucosyltransferase, MFUT-II,
RT and not MFUT-I or -III, is induced in a restricted region of the
RT digestive tract of germ-free mice by host-microbe interactions and
RT cycloheximide."
RL Biochim. Biophys. Acta 1487:275-285(2000).
DR EMBL; AF064752; AAC16887.1;
DR MGD; MGI:109374; Fut2.
DR InterPro; IPR002516; GT_11.
DR Pfam; PF01531; Glyco_transf_11.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 347 AA; 39215 MW; 616CCB11581C4179 CRC64;

Query Match 79.5%; Score 1471.5; DB 11; Length 347;
Best Local Similarity 79.0%; Pred. No. 7e-129;
Matches 275; Conservative 30; Mismatches 36; Indels 5; Gaps 2;

QY 1 MLVQMPSPFMAHILFVFTVSTIIHQRIYVKLQTLSEKELQAVQMSSPNAARDIMQ 59
Db 1 MASAOVPSFPLAHFLIFVFTVSTIIHQRIYVKLQTLSEKELQAVQMSSPNAARDIMQ 60
QY 60 ---LRGWTINAGRLGNQMGVATYALAKMNGRPAFIPAQMHSITLAPFRITLPLVLS 116
Db 61 SAKLOGIFTINSIGRLGNQMGVATYALAKMNGRPAFIPESMHNALAFIRISLPVLS 120
QY 117 ATASRIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 176
Db 121 DTARRIPWQNYHLNDWMEERYRHI-PGQYVREFTGYPGCSWTFYHHLRQELQOFTLHDHVR 179
QY 177 EEAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALDWRFRYS 236
Db 180 EEAQKFLRGLOVNGSRPGTFVGVHVRGDDYVHVPKVGWVADRRYLOQALDWRFRYS 239
QY 237 SLIFVVTNSGMWACRENIDTSHGDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 296
Db 240 SPVFTVTSNGMWCRENINSLGDDVVFAGDIEGSPAKDFALLTQCNHTIMTIGTFGIWA 299
QY 297 AYLTGSDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 344
Db 300 AYLAGGDTIYLANITLPSDFLKIFKPEAAFLPEWTGIAADLSPLKH 347

Search completed: May 27, 2003, 15:10:47
Job time : 37.6774 secs

```